Principles of Biomedical Ethics

EIGHTH EDITION

Tom L. Beauchamp

James F. Childress

PREFACE TO THE EIGHTH EDITION

Biomedical ethics, or bioethics, was a youthful field when the first edition of this book went to press in late 1977, now over forty years ago. The word *bioethics* was a recently coined term when, in the mid-1970s, we began as a team writing in this field and lecturing to health professionals on the subject of moral theory and principles. The field had virtually no literature that engaged moral theory and methodology. Massive changes have since occurred both in the field and in this book. We have tried to stay as close to the frontiers of this field as we could, even though the literature is now sufficiently extensive and rapidly expanding that it is difficult to keep abreast of new topics under discussion.

For those who have stayed with us through the previous editions of *Principles of Biomedical Ethics*, we express our gratitude for your critical and constructive suggestions—for us a constant source of information and insight, as well as inspiration. Substantial changes have appeared in all editions after the first, and this eighth and perhaps final edition is no exception. No new changes have been made in the book's basic structure, but the revisions are thoroughgoing in every chapter. We have attempted to sharpen our investigations, strengthen our arguments, address issues raised by critics, and both reference and assess new published material. As in previous editions, we have made changes in virtually every section and subsection of the book's ten chapters.

Our clarifications, additions, expansions, and responses to critics can be crisply summarized as follows:

Part I, *Moral Foundations*: In Chapter 1, "Moral Norms," we have clarified, augmented, and tightened our accounts of the common morality, universal morality, and how they differ from particular moralities. We have also clarified in this chapter and Chapter 10 the ways in which the four-principles framework is to be understood as a substantive framework of practical normative principles and a method of bioethics. We have had a major commitment to the virtues and moral character since our first edition. In Chapters 2 and 9 we have clarified and modestly expanded our discussion of the nature and importance of moral virtues, moral ideals, and moral excellence; and we have also revised our account of the lines that separate what is obligatory, what is beyond obligation, and what is virtuous. In Chapter 3, "Moral Status," we have revised our account of theories of moral status in several ways and revised our presentation in the section on "Guidelines Governing Moral Status: Putting Specification to Work." We also engage some moral problems that have emerged about the use of human-nonhuman chimeras in biomedical research. We there concentrate on whether functional integration of human neural cells in a nonhuman primate brain (and the brains of other species) would cause a morally significant change in the mind of the animal, and, if it did so, what the consequences should be for the moral status of the animal if it were born.

Part II, *Moral Principles*: The principles of basic importance for biomedical ethics are treated individually in Part II. In Chapter 4, "Respect for Autonomy," we have expanded our presentations in several sections including addition of an analysis of the distinction between the *justification* of informed consent requirements and the several *functions* served by the doctrine, institutions, and practices of informed consent. Also added is a significant clarification of our theory of intentional nondisclosure in clinical practice and research and the conditions under which intentional nondisclosure is justified. In Chapter 5, "Nonmaleficence," we have updated and deepened our constructive proposals about "Distinctions and Rules Governing Nontreatment," proper and improper uses of the best-interest standard, and the place of anticipated quality of life in decisions regarding seriously ill newborns and children. The sections on decisions about physician-assisted dying are updated and arguments adjusted in light of global developments, especially in North America (Canada and several US states). In Chapter 6, "Beneficence," we deepened our analysis of policies of expanded and continued access to investigational products in research as well as our discussions of the ethical value of, concerns about, and constraints on risk-benefit, cost-benefit, and cost-effectiveness analyses. In Chapter 7, "Justice," we updated and expanded the discussions of theories of justice, with restructured presentations of communitarian theories, capability theories, and well-being theories. Also updated are sections on problems of health insurance coverage, social implementation of the right to health care, and the right to a decent minimum of health care—as well as revised analyses of whether individuals forfeit this right through risky actions and what the fair opportunity rule requires by way of rectifying disparities in health care. Chapter 8, "Professional-Patient Relationships," has

expanded sections on "Veracity" and "Confidentiality," each of which incorporates new cases. The section on arguments for intentionally limiting communication of bad news has been updated. In particular, we have deepened our account of when physicians' decisions to use staged disclosures are ethically justified.

Part III, *Theory and Method*: Chapter 9, "Moral Theories," has an expanded section on "Virtue Theory" that fills out our account of the virtues introduced in Chapter 2 and furthers the application of our theory to biomedical ethics. We have also augmented and clarified the section on rights theory. Significant additions appear in the section on "The Rights of Incompetent, Disadvantaged, and Unidentified Members of Populations." In Chapter 10, "Method and Moral Justification," we have strengthened our critiques of theories of justification in what we call top-down models and casuistry. We have also expanded our accounts of common-morality theory, moral change, reflective equilibrium, considered judgments, and the ways in which our theory is committed to a global bioethics. Each of these parts has been recast to clarify and deepen our positions.

Finally, we want to correct some long-standing misinterpretations of our theory that have persisted over the forty years of editions of this book. Several critics have maintained that our book is committed to an American individualism in which the principle of respect for autonomy dominates all other moral principles and considerations. This interpretation of our book is profoundly mistaken. In a properly structured account of biomedical ethics, respect for autonomy has no distinctly American grounding and is not excessively individualistic or overriding. We do not emphasize individual rights to the neglect or exclusion of social responsibilities and communal goals. We do not now, and have never, treated the principle of respect for autonomy in the ways several of our critics allege. To the contrary, we have always argued that many competing moral considerations validly override this principle under certain conditions. Examples include the following: If our choices endanger public health, potentially harm innocent others, or require a scarce and unfunded resource, exercises of autonomy can justifiably be restricted by moral and legal considerations. The principle of respect for autonomy does not by itself determine what, on balance, a person ought to be free to do or what counts as a valid justification for constraining autonomy.

Our position is that it is a mistake in biomedical ethics to assign priority a priori to any basic principle over other basic principles—as if morality is hierarchically structured or as if we must value one moral norm over another without consideration of particular circumstances. The best strategy is to appreciate the contributions and the limits of various principles, virtues, and rights, which is the strategy we have embraced since the first edition and continue throughout this edition. A number of our critics have mistakenly maintained—without textual warrant—that our so-called principlism overlooks or even discounts the virtues. We have given a prominent place in our theory—since the first edition—to the virtues and their significant role in biomedical ethics. We maintain and further develop this commitment in the present edition.

Fortunately, we have always had a number of valuable—and often constructive—critics of our theories, especially John Arras, Edmund Pellegrino, Raanan Gillon, Al Jonsen, Stephen Toulmin, Michael Yesley, Franklin Miller, David DeGrazia, Ronald Lindsay, Carson Strong, John-Stewart Gordon, Oliver Rauprich, Jochen Vollmann, Rebecca Kukla, Henry Richardson, Peter Herissone-Kelly, Robert Baker, Robert Veatch, Tris Engelhardt, Robert "Skip" Nelson, and Neal W. Dickert. Our book owes a great deal to these critics and friends. We again wish to remember with great fondness and appreciation the late Dan Clouser, a wise man who seems to have been our first—and certainly one of our sternest—critics. We also acknowledge the penetrating criticisms of Clouser's friend, and ours, the late Bernard Gert, whose trenchant criticisms showed us the need for clarifications or modifications in our views. We also thank John Rawls for a lengthy conversation, shortly before his untimely death in 2002, about communitarian and egalitarian theories of justice that led to significant improvements in our chapter on justice.

We have continued to receive many helpful suggestions for improvements in our work from students, colleagues, health professionals, and teachers who use the book. Jim is particularly grateful to his University of Virginia colleagues: the late John Arras, already mentioned; Ruth Gaare Bernheim; Richard Bonnie; and the late John Fletcher for many illuminating discussions in team-taught courses and in other contexts. Discussions with many practicing physicians and nurses in the University of Virginia's Medical Center, on its Ethics Committee, and with faculty in the Center for Biomedical Ethics and Humanities have been very helpful. In addition, Jim thanks the faculty and graduate students of the Centre for the Advanced Study of Bioethics at the University of

Münster for gracious hospitality and vigorous and valuable conversation and debate, particularly about paternalism and autonomy, especially during extended visits in 2011 and 2016; Bettina Schöne-Seifert, Thomas Gutmann, and Michael Quante deserve special thanks. Jim also expresses his deep gratitude to Marcia Day Childress, his wife for the last twenty-two years, for many valuable suggestions along with loving and unstinting support throughout the preparation of the eighth edition as well as the preceding three editions.

Tom likewise wishes to thank his many colleagues in Georgetown University's Philosophy Department and Kennedy Institute of Ethics, as well as his colleagues in research at the Berman Institute of Bioethics of The Johns Hopkins University. Henry Richardson and Rebecca Kukla have been penetrating, as well as constructive, critics from whom several editions of this book have greatly benefited. Between the sixth and seventh editions, Tom benefited hugely from his work with colleagues at Johns Hopkins on an NIH grant to study the need to revise our understanding of the research–practice distinction: Ruth Faden, Nancy Kass, Peter Pronovost, Steven Goodman, and Sean Tunis. When one has colleagues this talented and well informed, multidisciplinary work is as invigorating as it is instructive.

Tom also wishes to express appreciation to five undergraduate research assistants: Patrick Connolly, Stacylyn Dewey, Traviss Cassidy, Kekenus Sidik, and Patrick Gordon. Their research in the literature, their editing of copy, and their help with previous indexes have made this book more comprehensive and readable. Likewise, Jim wishes to thank three superb research and teaching assistants, Matt Puffer, Travis Pickell, and Laura Alexander, for their helpful contributions. Other teaching assistants in a lecture course at the University of Virginia that used this book also made valuable suggestions.

We also acknowledge with due appreciation the support provided by the Kennedy Institute's library and information retrieval systems, which kept us in touch with new literature and reduced the burdens of library research. We owe a special debt of gratitude to Martina Darragh, who retired as the last chapter of this eighth edition was being completed. Martina gave us help when we thought no help could be found.

Retrospectively, we express our gratitude to Jeffrey House, our editor at Oxford University Press for the first thirty years of this book. Jeff encouraged us to write it before a single page was written, believed in it deeply, and saw it through all of its formative editions. He was an emulable editor. We also thank Robert Miller for efficiently facilitating the production of the recent editions of this book.

We dedicate this edition, just as we have dedicated each of the previous seven editions, to Georgia, Ruth, and Don. Georgia, Jim's beloved wife of thirty-five years, died in 1994, just after the fourth edition appeared. Our dedication honors her wonderful memory and her steadfast support for this project from its inception. Tom also acknowledges the love, devotion, and intellectual contribution to this book of his wife, Ruth Faden, who has been the deepest influence on his career in bioethics, and salutes Donald Seldin, a brilliant physician and an inspiration to Tom and to biomedical ethics since the early years of the field. Don passed away at age ninety-seven in 2018, when we were in the midst of preparing this eighth edition. He will be sorely missed, and never forgotten.

Washington, DC, and *Chilmark, MA* T.L.B.

Charlottesville, VA J.F.C.

January 2019

1

Moral Norms

In the last third of the twentieth century, major developments in the biological and health sciences and in biomedical technology strikingly challenged traditional professional ethics in much of clinical medicine, nursing, and biomedical and behavioral research.[1] Despite a remarkable continuity in medical ethics across millennia, the widely revered Hippocratic tradition could not adequately address modern concerns such as informed consent, privacy, access to health care, communal and public health responsibilities, and research involving human subjects. Professional ethics was also ill equipped to provide an adequate framework for public policy in a pluralistic society.

In this book, we acknowledge and draw from the great traditions of medical ethics,[2] but we also draw from philosophical reflections on morality. This approach helps us to examine and, where appropriate, challenge common assumptions in the biomedical sciences, health care, and public health.

Normative and Nonnormative Ethics

The term *ethics* needs attention before we turn to the meanings of *morality* and *professional ethics*. *Ethics* is a generic term covering several different ways of examining and interpreting the moral life. Some approaches to ethics are normative, others nonnormative.

Normative Ethics

General normative ethics addresses the question, "Which general moral norms should we use to guide and evaluate conduct, and why?" Ethical theories seek to identify and justify these norms, which are often referred to as principles, rules, rights, or virtues. In Chapter 9 we examine several types of general normative ethical theory and offer criteria for assessing them.

Many practical questions would remain unanswered even if a fully satisfactory general ethical theory were available. The term *practical ethics*, as used here, is synonymous with *applied ethics* and stands in contrast to *theoretical ethics*.[3] *Practical ethics* refers to the use of moral concepts and norms in deliberations about moral problems, practices, and policies in professions, institutions, and public policy. Often no direct movement from general norms, precedents, or theories to particular judgments is possible. General norms are usually only starting points for the development of more specific norms of conduct suitable for contexts such as clinical medicine and biomedical research. Throughout this book we address how to move from general norms to specific norms and particular judgments and from theory to practice.

Nonnormative Ethics

Two types of nonnormative ethics are distinguishable. The first is *descriptive ethics*, which is the factual investigation of moral beliefs and conduct. It often uses scientific techniques to study how people reason and act. For example, anthropologists, sociologists, psychologists, and historians determine which moral norms are expressed in professional practice, in professional codes, in institutional mission statements and rules, and in public policies. These researchers study phenomena such as surrogate decision making, treatment of the dying, the use of vulnerable populations in research, how consents are obtained from patients, and refusal of treatment by patients.

The second type of nonnormative ethics is *metaethics*, which involves analysis of the language, concepts, and methods of reasoning in normative ethics.[4] For example, metaethics addresses the meanings of terms such as

right, *obligation*, *virtue*, *justification*, *morality*, and *responsibility*. It is also concerned with moral epistemology (the theory of moral knowledge), the logic and patterns of moral reasoning and justification, and the nature and possibility of moral truth. Whether morality is objective or subjective, relative or nonrelative, and rational or nonrational are prominent questions in metaethics.

Descriptive ethics and metaethics are nonnormative because their objective is to establish what factually or conceptually *is* the case, not what ethically *ought to be* the case or what is ethically *valuable*. For example, in this book we often rely on reports in descriptive ethics when investigating the nature of professional conduct and codes of ethics, current forms of access to health care, and physician attitudes toward hastening the deaths of patients who have requested aid in dying. In these investigations we are interested in how such descriptive information assists us in determining which practices are morally justifiable as well as in resolving other normative issues.

THE COMMON MORALITY AS UNIVERSAL MORALITY

In its most familiar sense, the word *morality* (a broader term than *common morality*, which is discussed immediately below in the section on "The Nature of the Common Morality," and in more detail in Chapter 10, pp. 444–57) refers to norms about right and wrong human conduct that are widely shared and form a stable societal compact. As a social institution, morality encompasses many standards of conduct, including moral principles, rules, ideals, rights, and virtues. We learn about morality as we grow up, and we learn to distinguish between the part of morality that holds for everyone and moral norms that bind only members of specific communities or special groups such as physicians, nurses, or public health officials.

The Nature of the Common Morality

Some core tenets found in every acceptable particular morality are not relative to cultures, groups, or individuals. All persons living a moral life know and accept rules such as not to lie, not to steal others' property, not to punish innocent persons, not to kill or cause harm to others, to keep promises, and to respect the rights of others. All persons committed to morality do not doubt the relevance and importance of these universally valid rules. Violation of these norms is unethical and will generate feelings of remorse. The literature of biomedical ethics virtually never debates the merit or acceptability of these central moral norms. Debates do occur, however, about their precise meaning, scope, weight, and strength, often in regard to hard moral cases or current practices that merit careful scrutiny—such as when, if ever, physicians may justifiably withhold some aspects of a diagnostic finding from their patients.

We call the set of universal norms shared by all persons committed to morality *the common morality*. This morality is not merely *a* morality, in contrast to other moralities.[5] It is applicable to all persons in all places, and we appropriately judge all human conduct by its standards. The following norms are examples (far from a complete list) of generally binding *standards of action* (that is, rules of obligation) found in the common morality: (1) Do not kill, (2) Do not cause pain or suffering to others, (3) Prevent evil or harm from occurring, (4) Rescue persons in danger, (5) Tell the truth, (6) Nurture the young and dependent, (7) Keep your promises, (8) Do not steal, (9) Do not punish the innocent, and (10) Obey just laws.

The common morality also contains standards other than obligatory rules of conduct. Here are ten examples of *moral character traits*, or virtues, recognized in the common morality (again, not a complete list): (1) nonmalevolence (not harboring ill will toward others), (2) honesty, (3) integrity, (4) conscientiousness, (5) trustworthiness, (6) fidelity, (7) gratitude, (8) truthfulness, (9) lovingness, and (10) kindness. These virtues are universally admired traits of character.[6] A person is deficient in moral character if he or she lacks such traits. Negative traits that are the opposite of these virtues are *vices* (for example, malevolence, dishonesty, lack of integrity, cruelty, etc.). They are universally recognized as substantial moral defects. In this chapter we will say nothing further about moral character and the virtues and vices, because they are investigated in both Chapter 2 and a major section of Chapter 9 (pp. 31–45, 409–16).

In addition to the obligations and virtues just mentioned, the common morality supports *human rights* and endorses *moral ideals* such as charity and generosity. Philosophers debate whether one of these regions of the moral life—obligations, rights, or virtues—is more basic or more valuable than another, but in the common morality there is no reason to give primacy to any one area or type of norm. For example, human rights are not more basic than moral virtues in universal morality, and moral ideals should not be downgraded morally merely because people are not obligated to conform to them. An undue emphasis on any one of these areas or types of norms disregards the full scope of morality.[7]

Our account of universal morality in this chapter and Chapter 10 does not conceive of the common morality as ahistorical or a priori.[8] This problem in moral theory cannot be adequately engaged until our discussions in Chapter 10, and we offer now only three clarifications of our position: First, the common morality is a product of human experience and history and is a universally shared product. The origin of the norms of the common morality is no different in principle from the origin of the norms of a particular morality for a medical or other profession. Both are learned and transmitted in communities. The primary difference is that the common morality has authority in all communities, whereas particular moralities are authoritative only for specific groups. Second, we accept moral pluralism in *particular* moralities, as discussed later in this chapter (pp. 5–6), but we reject moral pluralism, understood as relativism, in the *common* morality. (See the section in Chapter 10 on "Moral Change" for further clarification.) No particular moral way of life qualifies as morally acceptable unless it conforms to the standards in the common morality. Third, the common morality comprises moral *beliefs* that all morally committed persons believe. It does not consist of timeless, detached standards of truth that exist independently of a history of moral beliefs. Likewise, every *theory* of the common morality has a history of development by the author(s) of the theory.

Ways to Examine the Common Morality

Various statements about or references to the common morality might be understood as normative, nonnormative, or possibly both. If the appeals are normative, the claim is that the common morality has normative force: It establishes moral standards for everyone, and violating these standards is unethical. If the references are nonnormative, the claim is that we can empirically study whether the common morality is present in all cultures. We accept both the normative force of the common morality and the objective of studying it empirically.

Some critics of our theory of the common morality (see Chapter 10) have asserted that scant anthropological or historical evidence supports the empirical hypothesis that a universal common morality exists.[9] Accordingly, they think we need to consider how good the evidence is both for and against the existence of a universal common morality. This problem is multifaceted and difficult to address, but in principle, scientific research could either confirm or falsify the hypothesis of a universal morality. It would be absurd to assert that all persons do in fact accept the norms of the common morality, because many amoral, immoral, or selectively moral persons do not care about or identify with its moral demands. Our hypothesis is that all persons *committed to morality* accept the standards in the common morality.

We explore this hypothesis about the empirical study of the common morality in Chapter 10 (pp. 449–52). Here we note only that when we claim that the normative judgments found in many parts of this book are derived from the common morality, we are not asserting that *our theory* of the common morality gets the common morality perfectly right or that it interprets or extends the common morality in just the right ways. There undoubtedly are dimensions of the common morality that we do not correctly capture or depict; and there are many parts of the common morality that we do not even address.[10] When we attempt to build on the common morality in this book by using it as a basis for critically examining problems of biomedical ethics, we do not mean to imply that our extensions can validly claim the authority of the common morality at every level of our interpretation of this morality.

PARTICULAR MORALITIES AS NONUNIVERSAL

We shift now from universal morality (the common morality) to particular moralities, which contain moral norms that are not shared by all cultures, groups, and individuals who are committed to morality.

The Nature of Particular Moralities

Whereas the common morality contains moral norms that are abstract, universal, and content-thin (such as "Tell the truth"), particular moralities present concrete, nonuniversal, and content-rich norms (such as "Make conscientious oral disclosures to, and obtain a written informed consent from, all human research subjects"). Particular moralities are distinguished by the specificity of their norms, but these norms are not morally justified if they violate norms in the common morality. Specific moralities include the many responsibilities, aspirations, ideals, sentiments, attitudes, and sensitivities found in diverse cultural traditions, religious traditions, professional practice, and institutional guides. Explication of the values in these moralities sometimes requires a special knowledge and may involve refinement by experts or scholars over centuries—as, for example, in the body of Jewish religious, legal, and moral norms in the Talmudic tradition; well-structured moral systems to provide methods for judgments and to adjudicate conflicts in Roman Catholic casuistry; and Islamic reliance on Shari'ah-based principles. Each tradition continues today to elaborate its commitments through the development of detailed, and hopefully coherent, systems of medical ethics. These elaborations are often derived from the common morality, not merely from the scriptures of a particular religious tradition.

Professional moralities, which include moral codes and standards of practice, are also particular moralities. They may legitimately vary from other moralities in the ways they handle certain conflicts of interest, research protocol reviews, advance directives, and similar matters. (See the next section below on "Professional and Public Moralities.") *Moral ideals* such as charitable goals and aspirations to rescue suffering persons in dangerous situations provide another instructive example of facets of particular moralities. By definition, moral ideals such as charitable beneficence are not morally *required* of all persons; indeed, they are not required of any person.[11] Persons who fail to fulfill even their own personal ideals cannot be blamed or criticized by others. These ideals may nonetheless be critically important features of personal or communal moralities. Examples are found in physicians' individual commitments or physician codes that call for assumption of a significant level of risk in circumstances of communicable disease. It is reasonable to presume that all morally committed persons share an admiration of and endorsement of moral ideals of generosity and service, and in this respect these ideals are part of shared moral beliefs in the common morality; they are universally *praiseworthy* even though not universally *required* or universally *practiced*. When such ideals are regarded by those who embrace them as obligations (as they are, for example, in some monastic traditions), the obligations are still parts of a particular morality, not of universal morality.

Persons who accept a particular morality sometimes presume that they can use this morality to speak with an authoritative moral voice for all persons. They operate under the false belief that their particular convictions have the authority of the common morality. These persons may have morally acceptable and even praiseworthy beliefs, but their particular beliefs do not bind other persons or communities. For example, persons who believe that scarce medical resources, such as transplantable organs, should be distributed by lottery rather than by medical need may have good moral reasons for their views, but they cannot claim that their views are supported by the common morality.

Professional and Public Moralities

Just as the common morality is accepted by all morally committed persons, most professions have, at least implicitly, a professional morality with standards of conduct that are generally acknowledged and encouraged by those in the profession who are serious about their moral responsibilities. In medicine, professional morality specifies general moral norms for the institutions and practices of medicine. Special roles and relationships in medicine derive from rules or traditions that other professions will likely not need or accept. As we argue in Chapters 4 and 8, rules of informed consent and medical confidentiality may not be serviceable or appropriate outside of medicine, nursing, biomedical research, and public health, but these rules are justified by general moral requirements of respecting the autonomy of persons and protecting them from harm.

Members of professions often adhere to moral guidelines such as rules prohibiting discrimination against colleagues on the basis of gender, race, religion, or national origin (some of these guidelines now have legal backing). In recent years formal codifications of and instruction in professional morality have increased through codes of medical and nursing ethics, codes of research ethics, corporate policies of bioethics, institutional guidelines governing conflict of interest, and the reports and recommendations of public commissions. Before we assess these guidelines, the nature of professions in general needs brief discussion.

In a classic work on the subject, Talcott Parsons defines a profession as "a cluster of occupational roles, that is, roles in which the incumbents perform certain functions valued in the society in general, and, by these activities, typically earn a living at a full-time job."[12] Under this definition, circus performers, exterminators, and garbage collectors are professionals. It is not surprising to find all such activities characterized as professions, inasmuch as the word *profession* has come, in common use, to mean almost any occupation by which a person earns a living. The once honorific sense of *profession* is now better reflected in the term *learned profession*, which assumes an extensive education in the arts, humanities, law, sciences, or technologies.

Professionals are usually distinguished by their specialized knowledge and training as well as by their commitment to provide important services or information to patients, clients, students, or consumers. Professions maintain self-regulating organizations that control entry into occupational roles by formally certifying that candidates have acquired the necessary knowledge and skills. In learned professions such as medicine, nursing, and public health, a professional's background knowledge is partly acquired through closely supervised training, and the professional is committed to providing a service to others.

Health care professions specify and enforce obligations for their members, thereby seeking to ensure that persons who enter into relationships with these professionals will find them competent and trustworthy.[13] The obligations that professions attempt to enforce are determined by an accepted role. These obligations comprise the "ethics" of the profession, although there may also be role-specific customs such as self-effacement that are not obligatory. Problems of professional ethics commonly arise either from conflicts over appropriate professional standards or conflicts between professional commitments and the commitments professionals have outside the profession.

Because traditional standards of professional morality are often vague, some professions codify their standards in detailed statements aimed at reducing vagueness and improving adherence. Their codes sometimes specify rules of etiquette in addition to rules of ethics. For example, a historically significant version of the code of the American Medical Association (AMA) dating from 1847 instructed physicians not to criticize fellow physicians who had previously been in charge of a case.[14] Such professional codes tend to foster and reinforce member identification with the prevailing values of the profession. These codes are beneficial when they effectively incorporate defensible moral norms, but some codes oversimplify moral requirements, make them indefensibly rigid, or make excessive and unwarranted claims about their completeness and authoritativeness. As a consequence, professionals may mistakenly suppose that they are satisfying all relevant moral requirements by scrupulously following the rules of the code, just as some people believe that they fully discharge their moral obligations when they meet all relevant legal requirements.

We can and should ask whether the codes specific to areas of science, medicine, nursing, health care, and public health are coherent, defensible, and comprehensive within their domain. Historically, few codes had much to say about the implications of several pivotal moral principles and rules such as veracity, respect for autonomy, and social justice that have been the subjects of intense discussion in recent biomedical ethics. From ancient medicine to the present, physicians have generated codes without determining their acceptability to patients and the public. These codes have rarely appealed to general ethical standards or to a source of moral authority beyond the traditions and judgments of physicians themselves.[15] The articulation of such professional norms has often served more to protect the profession's interests than to offer a broad and impartial moral viewpoint or to address issues of importance to patients and society.[16]

Psychiatrist Jay Katz poignantly expressed reservations about traditional principles and codes of medical ethics. Initially inspired by his outrage over the fate of Holocaust victims at the hands of German physicians, Katz

became convinced that a professional ethics that reaches beyond traditional codes is indispensable:

> As I became increasingly involved in the world of law, I learned much that was new to me from my colleagues and students about such complex issues as the right to self-determination and privacy and the extent of the authority of governmental, professional, and other institutions to intrude into private life. … These issues … had rarely been discussed in my medical education. Instead it had been all too uncritically assumed that they could be resolved by fidelity to such undefined principles as *primum non nocere* ["First, do no harm"] or to visionary codes of ethics.[17]

The Regulation and Oversight of Professional Conduct

Additional moral direction for health professionals and scientists comes through the public policy process, which includes regulations and guidelines promulgated by governmental bodies. The term *public policy* refers to a set of normative, enforceable guidelines adopted by an official public body, such as an agency of government or a legislature, to govern a particular area of conduct. The policies of corporations, hospitals, trade groups, and professional societies are private, not public, even if these bodies are regulated to some degree by public policies and sometimes have an impact on public policy.

A close connection exists between law and public policy: All laws constitute public policies, but not all public policies are, in the conventional sense, laws. In contrast to laws, public policies need not be explicitly formulated or codified. For example, an official who decides not to fund a newly recommended government program with no prior history of funding is formulating a public policy. Decisions not to act, as well as decisions to act, can constitute policies.

Policies such as those that fund health care for the indigent or that protect subjects of biomedical research regularly incorporate moral considerations. Moral analysis is part of good policy formation, not merely a method for evaluating existing policy. Efforts to protect the rights of patients and research subjects are instructive examples. Over the past few decades many governments have created national commissions, national review committees, advisory committees, and councils to formulate guidelines for research involving human subjects, for the distribution of health care, and for addressing moral mistakes made in the health professions. Morally informed policies have guided decision making about other areas of practice as well. The relevance of bioethics to public policy is now recognized in most countries, some of which have influential standing bioethics committees.[18]

Many courts have developed case law that sets standards for science, medicine, and health care. Legal decisions often express communal moral norms and stimulate ethical reflection that over time alters those norms. For example, the lines of court decisions in many countries about how dying patients may be or must be treated have constituted nascent traditions of moral reflection that have been influenced by, and in turn have influenced, literature in biomedical ethics on topics such as when artificial devices that sustain life may be withdrawn, whether medically administered nutrition and hydration is a medical treatment that may be discontinued, and whether physicians may be actively involved in hastening a patient's death at the patient's request.

Policy formation and criticism generally involve more specific moral judgments than the judgments found in general ethical theories, principles, and rules.[19] Public policy is often formulated in contexts that are marked by profound social disagreements, uncertainties, and differing interpretations of history. No body of abstract moral principles and rules can fix policy in such circumstances, because abstract norms do not contain enough specific information to provide direct and discerning guidance. The implementation of moral principles and rules, through specification and balancing, must take into account factors such as feasibility, efficiency, cultural pluralism, political procedures, pertinent legal requirements, uncertainty about risk, and noncompliance by patients. Moral principles and rules provide a normative structure for policy formation and evaluation, but policies are also shaped by empirical data and information generated in fields such as medicine, nursing, public health, veterinary science, economics, law, biotechnology, and psychology.

When using moral norms to formulate or criticize public policies, one cannot move with assurance from a judgment that an *act* is morally right (or wrong) to a judgment that a corresponding *law* or *policy* is morally right (or wrong). Considerations such as the symbolic value of law and the costs of a publicly funded program and its enforcement often may have substantial importance for law and policy. The judgment that an act is morally wrong does not entail the judgment that the government should prohibit it or refuse to allocate funds to support it. For example, one can argue without any inconsistency that sterilization and abortion are morally wrong but that the law should not prohibit them, because they are fundamentally matters of personal choice beyond the legitimate reach of government—or, alternatively, because many persons would seek dangerous and unsanitary procedures from unlicensed practitioners. Similarly, the judgment that an act is morally acceptable does not imply that the law should permit it. For example, the belief that euthanasia is morally justified for some terminally ill infants who face uncontrollable pain and suffering is consistent with the belief that the government should legally prohibit such euthanasia on grounds that it would not be possible to control abuses if it were legalized.

We are not defending any of these moral judgments. We are maintaining only that the connections between moral norms and judgments about policy or law are complicated and that a judgment about the morality of particular actions does not entail a comparable judgment about law or policy.

MORAL DILEMMAS

Common to all forms of practical ethics is reasoning through difficult cases, some of which constitute dilemmas. This is a familiar feature of decision making in morality, law, and public policy. Consider a classic case[20] in which judges on the California Supreme Court had to reach a decision about the legal force and limits of medical confidentiality. A man had killed a woman after confiding to a therapist his intention to do so. The therapist had attempted unsuccessfully to have the man committed but, in accordance with his duty of medical confidentiality to the patient, did not communicate the threat to the woman when the commitment attempt failed.

The majority opinion of the court held that "When a therapist determines, or pursuant to the standards of his profession should determine, that his patient presents a serious danger of violence to another, he incurs an obligation to use reasonable care to protect the intended victim against such danger." This obligation extends to notifying the police and also to warning the intended victim. The justices in the majority opinion argued that therapists generally ought to observe the rule of medical confidentiality, but that the rule must yield in this case to the "public interest in safety from violent assault." These justices recognized that rules of professional ethics have substantial public value, but they held that matters of greater importance, such as protecting persons against violent assault, can override these rules.

In a minority opinion, a judge disagreed and argued that doctors violate patients' rights if they fail to observe standard rules of confidentiality. If it were to become common practice to break these rules, he reasoned, the fiduciary nature of the relationship between physicians and patients would erode. Persons who are mentally ill would refrain from seeking aid or divulging critical information because of the loss of trust that is essential for effective treatment.

This case presents moral and legal dilemmas in which the judges cite relevant reasons to support their conflicting judgments.[21] Moral dilemmas are circumstances in which moral obligations demand or appear to demand that a person adopt each of two (or more) alternative but incompatible actions, such that the person cannot perform all the required actions. These dilemmas occur in at least two forms.[22] (1) Some evidence or argument indicates that an act is morally permissible and some evidence or argument indicates that it is morally wrong, but the evidence or strength of argument on both sides is inconclusive. Abortion, for example, may present a terrible dilemma for women who see the evidence in this way. (2) An agent believes that, on moral grounds, he or she is obligated to perform two or more mutually exclusive actions. In a moral dilemma of this form, one or more moral norms obligate an agent to do *x* and one or more moral norms obligate the agent to do *y*, but the agent cannot do both in the circumstance. The reasons behind alternatives *x* and *y* are weighty and neither set of reasons is overriding. If one acts on either set of reasons, one's actions will be morally acceptable

in some respects and morally unacceptable in others. The withdrawal of life-prolonging therapies from patients suffering from a wakeful unconscious state (formerly called a persistent, continuing, or continuous vegetative state) is sometimes regarded as an instance of this second form of dilemma.

Popular literature, novels, and films often illustrate how conflicting moral principles and rules create difficult dilemmas. For example, an impoverished person who steals from a grocery store to save a family from starvation confronts such a dilemma. The only way to comply with one obligation is to contravene another obligation. Some obligation must be overridden or compromised no matter which course is chosen. From the perspective we defend, it is confusing to say that we are obligated to perform both actions in these dilemmatic circumstances. Instead, we should discharge the obligation that we judge to override what we would have been firmly obligated to perform were it not for the conflict.

Conflicts between moral requirements and self-interest sometimes create a *practical* dilemma, but not, strictly speaking, a *moral* dilemma. If moral reasons compete with nonmoral reasons, such as self-interest, questions about priority can still arise even though no moral dilemma is present. When a moral reason conflicts with a personal reason, the moral reason is not always overriding. If, for example, a physician must choose between saving his or her own life or that of a patient, in a situation of extreme scarcity of available drugs, the moral obligation to take care of the patient may not be overriding.

Some moral philosophers and theologians have argued that although many practical dilemmas involving moral reasons exist, no irresolvable moral dilemmas exist. They do not deny that agents experience moral perplexity or conflict in difficult cases. However, they claim that the purpose of a moral theory is to provide a principled procedure for resolving deep conflicts. Some philosophers have defended this conclusion because they accept one supreme moral value as overriding all other conflicting values (moral and nonmoral) and because they regard it as incoherent to allow contradictory obligations in a properly structured moral theory. The only *ought*, they maintain, is the one generated by the supreme value.[23] (We examine such theories, including both utilitarian and Kantian theories, in Chapter 9.)

In contrast to the account of moral obligation offered by these theories, we maintain throughout this book that various moral principles, rules, and rights can and do conflict in the moral life. These conflicts sometimes produce irresolvable moral dilemmas. When forced to a choice, we may "resolve" the situation by choosing one option over another, but we also may believe that neither option is morally preferable. A physician with a limited supply of medicine may have to choose to save the life of one patient rather than another and still find his or her moral dilemma irresolvable. Explicit acknowledgment of such dilemmas helps deflate unwarranted expectations about what moral principles and theories can do. Although we find ways of reasoning about what we should do, we may not be able to reach a reasoned resolution in many instances. In some cases the dilemma becomes more difficult and remains unresolved even after the most careful reflection.

A Framework of Moral Principles

Moral norms central to biomedical ethics rely on the common morality, but they do not exhaust the common morality. Some types of basic moral norms are treated in this section, especially principles, rules, and rights. The virtues are the subject of Chapter 2, and the principles of primary importance for biomedical ethics are treated individually in Part II of this book. Most classical ethical theories accept these norms in some form, and traditional medical codes incorporate or presuppose at least some of them.

Principles

The set of pivotal moral principles defended in this book functions as an analytical framework of general norms derived from the common morality that form a suitable starting point for reflection on moral problems in biomedical ethics.[24] These principles are general guidelines for the formulation of more specific rules. In Chapters 4 through 7 we defend four clusters of moral principles: (1) *respect for autonomy* (a norm of respecting and supporting autonomous decisions), (2) *nonmaleficence* (a norm of avoiding the causation of harm), (3)

beneficence (a group of norms pertaining to relieving, lessening, or preventing harm and providing benefits and balancing benefits against risks and costs), and (4) *justice* (a cluster of norms for fairly distributing benefits, risks, and costs).

Nonmaleficence and beneficence have played central roles in the history of medical ethics. By contrast, respect for autonomy and justice were neglected in traditional medical ethics and have risen to prominence in this field only recently. In 1803, British physician Thomas Percival published *Medical Ethics*, the first comprehensive account of medical ethics in the long history of the subject. This book served as the backbone of British medical ethics and as the prototype for the American Medical Association's first code of ethics in 1847. Percival argued, using somewhat different language, that nonmaleficence and beneficence fix the physician's primary obligations and triumph over the patient's preferences and decision-making rights in circumstances of conflict.[25] Percival understated the critically important place of principles of respect for autonomy and distributive justice for physician conduct, but, in fairness to him, these considerations are now prominent in discussions of ethics in medicine in a way they were not when he wrote *Medical Ethics*.

That these four clusters of moral principles are central to biomedical ethics is a conclusion the authors of this work have reached by examining *considered moral judgments* and the *coherence of moral beliefs*, two notions analyzed in Chapter 10. The selection of these four principles, rather than some other clusters of principles, does not receive an argued defense in Chapters 1 through 3. However, in Chapters 4 through 7, we defend the vital role of each principle in biomedical ethics.

Rules

The framework of moral norms in this book encompasses several types of normative guidance, most notably principles, rules, rights, and virtues. Principles are more comprehensive and less specific than rules, but we draw only a loose distinction between them. Both are norms of obligation, but rules are more specific in content and more restricted in scope. Principles do not function as precise guides in each circumstance in the way that more detailed rules and judgments do. Principles and rules of obligation have correlative rights and often corresponding virtues. (See the discussion of rights in Chapter 9 and of virtues in Chapter 2.)

We defend several types of rules, the most important being substantive rules, authority rules, and procedural rules.

Substantive rules. Rules of truth telling, confidentiality, privacy, forgoing treatment, informed consent, and rationing health care provide more specific guides to action than do abstract principles. An example of a rule that sharpens the requirements of the principle of respect for autonomy in certain contexts is "Follow an incompetent patient's advance directive whenever it is clear and relevant." To indicate how this rule *specifies* the principle of respect for autonomy, it needs to be stated in full as "Respect the autonomy of incompetent patients by following all clear and relevant formulations in their advance directives." This specification shows how the initial norm of respect for autonomy endures even while becoming specified. (See the subsection "Specifying Principles and Rules" in the next section of this chapter.)

Authority rules. We also defend rules of decisional authority—that is, rules regarding who may and should make decisions and perform actions. For example, *rules of surrogate authority* determine who should serve as surrogate agents when making decisions for incompetent persons; *rules of professional authority* determine who in professional ranks should make decisions to accept or to override a patient's decisions; and *rules of distributional authority* determine who should make decisions about allocating scarce medical resources such as new and expensive medical technologies.

Authority rules do not delineate substantive standards or criteria for making decisions. However, authority rules and substantive rules interact in some situations. For instance, authority rules are justified, in part, by how well particular authorities can be expected to respect and comply with substantive rules and principles.

Procedural rules. We also defend rules that establish procedures to be followed. Procedures for determining eligibility for organ transplantation and procedures for reporting grievances to higher authorities are typical examples. We often resort to procedural rules when we run out of substantive rules and when authority rules are incomplete or inconclusive. For example, if substantive or authority rules are inadequate to determine which patients should receive scarce medical resources, a resort to procedural rules such as queuing and lottery may be justifiable.[26]

CONFLICTING MORAL NORMS

Prima Facie Obligations and Rights

Principles, rules, obligations, and rights are not rigid or absolute standards that allow no compromise. Although "a person of principle" is sometimes depicted as strict and unyielding, principles must be balanced and specified so they can function practically. It is no objection to moral norms that, in some circumstances, they can be justifiably overridden by other norms with which they conflict. All general moral norms are justifiably overridden in some circumstances. For example, we might justifiably not tell the truth to prevent someone from killing another person; and we might justifiably disclose confidential information about a person to protect the rights of another person.

Actions that harm individuals, cause basic needs to go unmet, or limit liberties are often said to be either wrong *prima facie* (i.e., wrongness is upheld unless the act is justifiable because of norms that are more stringent in the circumstances) or wrong *pro tanto* (i.e., wrong to a certain extent or wrong unless there is a compelling justification)—which is to say that the action is wrong in the absence of other moral considerations that supply a compelling justification.[27] Compelling justifications are sometimes available. For example, in circumstances of a severe swine flu pandemic, the forced confinement of persons through isolation and quarantine orders might be justified. Here a justifiable infringement of liberty rights occurs.

W. D. Ross's distinction between *prima facie* and *actual* obligations clarifies this idea. A *prima facie* obligation must be fulfilled unless it conflicts with an equal or stronger obligation. Likewise, a *prima facie* right (here we extend Ross's theory) must prevail unless it conflicts with an equal or stronger right (or conflicts with some other morally compelling alternative). Obligations and rights always constrain us unless a competing moral obligation or right can be shown to be overriding in a particular circumstance. As Ross put it, agents can determine their *actual* obligations in situations of conflict by examining the respective weights of the competing prima facie obligations. What agents ought to do is determined by what they ought to do all things considered.[28]

Imagine that a psychiatrist has confidential medical information about a patient who also happens to be an employee in the hospital where the psychiatrist practices. The employee seeks advancement in a stress-filled position, but the psychiatrist has good reason to believe that this advancement would be devastating for both the employee and the hospital. The psychiatrist has several prima facie duties in these circumstances, including those of confidentiality, nonmaleficence, beneficence, and respect for autonomy. Should the psychiatrist break confidence in this circumstance to meet these other duties? Could the psychiatrist make "confidential" disclosures to a hospital administrator and not to the personnel office? Addressing such questions through moral deliberation and justification is required to establish an agent's actual duty in the face of the conflicting prima facie duties.

These matters are more complicated than Ross suggests, particularly when rights come into conflict. We may need to develop a structured moral system or set of guidelines in which (1) some rights in a certain class of rights (for example, rights of individuals while alive to decide whether to donate their tissues and organs after death) have a fixed priority over others in another class of rights (for example, rights of family members to make decisions about the donation of their deceased relatives' tissues and organs) and (2) morally compelling social objectives such as gathering information in biomedical research can generally be overridden by basic human rights such as the right to give an informed consent or refusal.

No moral theory or professional code of ethics has successfully presented a system of moral rules free of conflicts and exceptions, but this observation should not generate either skepticism or alarm about ethical reflection, argument, and theory. The distinction between prima facie and actual obligations conforms closely to our experience as moral agents and provides indispensable categories for biomedical ethics. Almost daily we confront situations that force us to choose among conflicting values in our personal lives. For example, a person's financial situation might require that he or she choose between buying books for school and buying a train ticket to see friends. Not having the books will be an inconvenience and a loss, whereas not visiting with friends will disappoint the friends. Such choices do not come effortlessly, but we are usually able to think through the alternatives, deliberate, and reach a conclusion.

Moral Regret and Residual Obligation

An agent who determines that a particular act is the best one to perform in a situation of conflicting obligations may still not be able to discharge all aspects of moral obligation by performing that act. Even the morally best action in the circumstances may still be regrettable and may leave a moral residue, also called a moral trace.[29] Regret and residue over what is not done can arise even if the right action is clear and uncontested.

This point is about continuing obligation, not merely about feelings of regret and residue. Moral residue occurs because a prima facie obligation does not simply disappear when overridden. Often we have residual obligations because the obligations we were unable to discharge create new obligations. We may feel deep regret and a sting of conscience, but we also realize that we have a duty to bring closure to the situation.[30] We can sometimes make up for not fulfilling an obligation in one or more of several ways. For example, we may be able to notify persons in advance that we will not be able to keep a promise; we may be able to apologize in a way that heals a relationship; we may be able to change circumstances so that the conflict does not occur again; and we may be able to provide adequate compensation.

Specifying Principles and Rules

The four clusters of principles we present in this book do not by themselves constitute a general ethical theory. They provide only a framework of norms with which to get started in biomedical ethics. These principles must be specified in order to achieve more concrete guidance. Specification is a process of reducing the indeterminacy of abstract norms and generating rules with action-guiding content.[31] For example, without further specification, "do no harm" is too bare for thinking through problems such as whether it is permissible to hasten the death of a terminally ill patient.

Specification is not a process of producing or defending general norms such as those in the common morality; it assumes that the relevant general norms are available. Specifying the norms with which one starts—whether those in the common morality or norms previously specified—is accomplished by narrowing the scope of the norms, not by explaining what the general norms mean. We narrow the scope, as Henry Richardson puts it, by "spelling out where, when, why, how, by what means, to whom, or by whom the action is to be done or avoided."[32] For example, the norm that we are obligated to "respect the autonomy of persons" cannot, unless specified, handle complicated problems in clinical medicine and research involving human subjects. A definition of "respect for autonomy" (e.g., as "allowing competent persons to exercise their liberty rights") clarifies one's meaning in using the norm, but it does not narrow the scope of the general norm or render it more specific in guiding actions.

Specification adds content. For example, as noted previously, one possible specification of "Respect the autonomy of patients" is "Respect the autonomy of competent patients by following their advance directives when they become incompetent." This specification will work well in some medical contexts, but it will confront limits in others, where additional specification will be needed. Progressive specification can continue indefinitely, but to qualify all along the way as a specification some transparent connection must be maintained to the initial general norm that gives moral authority to the resulting string of specifications. This process is a

prime way in which general principles become practical instruments for moral reasoning; and it also helps explain why the four-principles approach is not merely an abstract theory limited to four general principles.[33]

An example of specification arises when psychiatrists conduct forensic evaluations of patients in a legal context. Psychiatrists cannot always obtain an informed consent, but they then risk violating their obligations to respect autonomy, a central imperative of medical ethics. A specification aimed at handling this problem is "Respect the autonomy of persons who are the subjects of forensic evaluations, where consent is not legally required, by disclosing to the evaluee the nature and purpose of the evaluation." We do not claim that this formulation is the best specification, but it approximates the provision recommended in the "Ethical Guidelines for the Practice of Forensic Psychiatry" of the American Academy of Psychiatry and the Law.[34] This specification attempts to guide forensic psychiatrists in discharging their diverse moral obligations.

Another example of specification derives from the oft-cited rule "Doctors should put their patients' interests first." In some countries patients are able to receive the best treatment available only if their physicians falsify information on insurance forms. The rule of patient priority does not imply that a physician should act illegally by lying or distorting the description of a patient's problem on an insurance form. Rules against deception, on the one hand, and for patient priority, on the other, are not categorical imperatives. When they conflict, we need some form of specification to know what we can and cannot do.

A survey of practicing physicians' attitudes toward deception illustrates how some physicians reconcile their dual commitment to patients and to nondeception. Dennis H. Novack and several colleagues used a questionnaire to obtain physicians' responses to difficult ethical problems that potentially could be resolved by use of deception. In one scenario, a physician recommends an annual screening mammography for a fifty-two-year-old woman who protests that her insurance company will not cover the test. The insurance company will cover the costs if the physician states (deceptively in this scenario) that the reason is "rule out cancer" rather than "screening mammography." The insurance company understands "rule out cancer" to apply only if there is a breast mass or other objective clinical evidence of the possibility of cancer, neither of which is present in this case. Almost 70% of the physicians responding to this survey indicated that they would state that they were seeking to "rule out cancer," and 85% of this group (85% of the 70%) insisted that their act would not involve "deception."[35]

These physicians' decisions are rudimentary attempts to specify the rule that "Doctors should put their patients' interests first." Some doctors seem to think that it is properly specified as follows: "Doctors should put their patients' interests first by withholding information from or misleading someone who has no *right* to that information, including an insurance company that, through unjust policies of coverage, forfeits its right to accurate information." In addition, most physicians in the study apparently did not operate with the definition of "deception" favored by the researchers, which is "to deceive is to make another believe what is not true, to mislead." Some physicians apparently believed that "deception" occurs when one person unjustifiably misleads another, and that it was justifiable to mislead the insurance company in these circumstances. It appears that these physicians would not agree on how to specify rules against deception or rules assigning priority to patients' interests.

All moral rules are, in principle, subject to specification. All will need additional content, because, as Richardson puts it, "the complexity of the moral phenomena always outruns our ability to capture them in general norms."[36] Many already specified rules will need further specification to handle new circumstances of conflict. These conclusions are connected to our earlier discussion of particular moralities. Different persons and groups will offer conflicting specifications, potentially creating multiple particular moralities. In any problematic case, competing specifications are likely to be offered by reasonable and fair-minded parties, all of whom are committed to the common morality.

To say that a problem or conflict is resolved or dissolved by specification is to say that norms have been made sufficiently determinate in content that, when cases fall under them, we know what must be done. Obviously some proposed specifications will fail to provide the most adequate or justified resolution. When competing specifications emerge, the proposed specifications should be based on deliberative processes of reasoning.

Specification as a method can be connected to a model of justification that will support some specifications and not others, as we argue in Chapter 10 (pp. 456–57).

Some specified norms are virtually absolute and need no further specification, though they are rare. Examples include prohibitions of cruelty that involve unnecessary infliction of pain and suffering.[37] "Do not rape" is a comparable example. More interesting are norms that are intentionally formulated with the goal of including all legitimate exceptions. An example is "Always obtain oral or written informed consent for medical interventions with competent patients, *except* in emergencies, in forensic examinations, in low-risk situations, or when patients have waived their right to adequate information." This norm needs further interpretation, including an analysis of what constitutes an informed consent, an emergency, a waiver, a forensic examination, and a low risk. This rule would be absolute if all legitimate exceptions had been successfully incorporated into its formulation, but such rules are rare. In light of the range of possibilities for contingent conflicts among rules, even the firmest and most detailed rules are likely to encounter exceptive cases.

Weighing and Balancing

Principles, rules, obligations, and rights often must be balanced in circumstances of contingent conflict. Does balancing differ from specification, or are they identical?

The process of weighing and balancing. Balancing occurs in the process of reasoning about which moral norms should prevail when two or more of them come into conflict. Balancing is concerned with the relative weights and strengths of different moral norms, whereas specification is concerned primarily with their range and scope, that is, their reach when narrowing the scope of pre-existing general norms (while adding content). Balancing consists of deliberation and judgment about these weights and strengths. It is well suited for reaching judgments in *particular cases*, whereas specification is especially useful for developing more *specific policies* from already accepted general norms.

The metaphor of larger and smaller weights moving a scale up and down has often been invoked to depict the balancing process, but this metaphor can obscure what happens in balancing. Justified acts of balancing are supported by good reasons. They need not rest merely on intuition or feeling, although intuitive balancing is one form of balancing. Suppose a physician encounters an emergency case that would require her to extend an already long day, making her unable to keep a promise to take her son to the local library. She engages in a process of deliberation that leads her to consider how urgently her son needs to get to the library, whether they could go to the library later, whether another physician could handle the emergency case, and the like. If she determines to stay deep into the night with the patient, she has judged this obligation to be overriding because she has found a good and sufficient reason for her action. The reason might be that a life hangs in the balance and she alone may have the knowledge to deal adequately with the circumstances. Canceling her evening with her son, distressing as it will be, could be justified by the significance of her reasons for doing what she does.

One way of approaching balancing merges it with specification. In our example, the physician's reasons can be generalized to similar cases: "If a patient's life hangs in the balance and the attending physician alone has the knowledge to deal adequately with the full array of the circumstances, then the physician's conflicting domestic obligations must yield." Even if we do not always state the way we balance considerations in the form of a specification, might not all deliberative judgments be made to conform to this model? If so, then deliberative balancing would be nothing but deliberative specification.

The goal of merging specification and balancing is appealing, but it is not well-suited to handle all situations in which balancing occurs. Specification requires that a moral agent extend norms by both narrowing their scope and generalizing to relevantly similar circumstances. Accordingly, "Respect the autonomy of competent patients when they become incompetent by following their advance directives" is a rule suited for all incompetent patients with advance directives. However, the responses of caring moral agents, such as physicians and nurses, are often highly specific to the needs of *this* patient or *this* family in *this* particular circumstance. Numerous considerations must be weighed and balanced, and any generalizations that could be formed might not hold even in remarkably similar cases.

Generalizations conceived as policies might even be dangerous. For example, cases in which risk of harm and burden are involved for a patient are often circumstances unlikely to be decided by expressing, by a rule, how much risk is allowable or how heavy the burden can be to secure a certain stated benefit. After levels of risk and burden are determined, these considerations must be balanced with the likelihood of the success of a procedure, the uncertainties involved, whether an adequately informed consent can be obtained, whether the family has a role to play, and the like. In this way, balancing allows for a due consideration of all the factors bearing on a complex particular circumstance, including all relevant moral norms.

Consider the following discussion with a young woman who has just been told that she is HIV-infected, as recorded by physician Timothy Quill and nurse Penelope Townsend:[38]

> PATIENT: Please don't tell me that. Oh my God. Oh my children. Oh Lord have mercy. Oh God, why did He do this to me? …
>
> DR. QUILL: First thing we have to do is learn as much as we can about it, because right now you are okay.
>
> PATIENT: I don't even have a future. Everything I know is that you gonna die anytime. What is there to do? What if I'm a walking time bomb? People will be scared to even touch me or say anything to me.
>
> DR. QUILL: No, that's not so.
>
> PATIENT: Yes they will, 'cause I feel that way …
>
> DR. QUILL: There is a future for you …
>
> PATIENT: Okay, all right. I'm so scared. I don't want to die. I don't want to die, Dr. Quill, not yet. I know I got to die, but I don't want to die.
>
> DR. QUILL: We've got to think about a couple of things.

Quill and Townsend work to calm down and reassure this patient, while engaging sympathetically with her feelings and conveying the presence of knowledgeable medical authorities. Their emotional investment in the patient's feelings is joined with a detached evaluation of the patient. Too much compassion and emotional investment may doom the task at hand; too much detachment will be cold and may destroy the patient's trust and hope. A balance in the sense of a right mixture between engagement and detachment must be found.

Quill and Townsend could try to specify norms of respect and beneficence to indicate how caring physicians and nurses should respond to patients who are desperately upset. However, specification will ring hollow and will not be sufficiently nuanced to provide practical guidance for this patient and certainly not for all desperately upset patients. Each encounter calls for a response inadequately captured by general principles and rules and their specifications. Behavior that is a caring response for one desperate patient may intrude on privacy or irritate another desperate patient. A physician may, for example, find it appropriate to touch or caress a patient, while appreciating that such behavior would be entirely inappropriate for another patient in a similar circumstance.

How physicians and nurses balance different moral considerations often involves sympathetic insight, humane responsiveness, and the practical wisdom of discerning a particular patient's circumstance and needs.[39] Balancing is often a more complex set of activities than those involved in a straightforward case of balancing two conflicting principles or rules. Considerations of trust, compassion, objective assessment, caring responsiveness, reassurance, and the like may all be involved in the process of balancing.

In many clinical contexts it may be hopelessly complicated and unproductive to engage in specification. For example, in cases of balancing harms of treatment against the benefits of treatment for incompetent patients, the

cases are often so exceptional that it is perilous to generalize a conclusion that would reach out to other cases. These problems are sometimes further complicated by disagreements among family members about what constitutes a benefit, poor decisions and indecision by a marginally competent patient, limitations of time and resources, and the like.[40]

We do not suggest that balancing is inescapably intuitive and unreflective. Instead, we propose a model of moral judgment that focuses on how balancing and judgment occur through practical astuteness, discriminating intelligence, and sympathetic responsiveness that are not reducible to the specification of norms. The capacity to balance many moral considerations is connected to what we discuss in Chapter 2 as capacities of moral character. Capacities in the form of virtues of compassion, attentiveness, discernment, caring, and kindness are integral to the way wise moral agents balance diverse, sometimes competing, moral considerations.

Practicability supplies another reason to support the conclusion that the model of specification needs supplementation by the model of balancing. Progressive specification covering all areas of the moral life would eventually mushroom into a body of norms so bulky that the normative system would become unwieldy. A scheme of comprehensive specification would constitute a package of potentially hundreds, thousands, or millions of rules, each suited to a narrow range of conduct. In the model of specification, every type of action in a circumstance of the contingent conflict of norms would be covered by a rule, but the formulation of rules for every circumstance of contingent conflict would be a body of rules too cumbersome to be helpful.

Conditions that constrain balancing. To allay concerns that the model of balancing is too intuitive or too open-ended and lacks a commitment to firm principles and rigorous reasoning, we propose six conditions that should help reduce intuition, partiality, and arbitrariness. These conditions must be met to justify infringing one prima facie norm in order to adhere to another.

1. 1. Good reasons are offered to act on the overriding norm rather than the infringed norm.
2. 2. The moral objective justifying the infringement has a realistic prospect of achievement.
3. 3. No morally preferable alternative actions are available.[41]
4. 4. The lowest level of infringement, commensurate with achieving the primary goal of the action, has been selected.
5. 5. All negative effects of the infringement have been minimized.
6. 6. All affected parties have been treated impartially.

Although some of these conditions are obvious and noncontroversial, some are often overlooked in moral deliberation and would lead to different conclusions were they observed. For example, some decisions to use futile life-extending technologies over the objections of patients or their surrogates violate condition 2 by endorsing actions in which no realistic prospect exists of achieving the goals of a proposed intervention. Typically, these decisions are made when health professionals regard the intervention as legally required, but in some cases the standard invoked is merely traditional or deeply entrenched.

Condition 3 is more commonly violated. Actions are regularly performed in some settings without serious consideration of alternative actions that might be performed. As a result, agents fail to identify a morally preferable alternative. For example, in animal care and use committees a common conflict involves the obligation to approve a good scientific protocol and the obligation to protect animals against unnecessary suffering. A protocol may be approved if it proposes a standard form of anesthesia. However, standard forms of anesthesia are not always the best way to protect the animal, and further inquiry is needed to determine the best anesthetic for the particular interventions proposed. In our schema of conditions, it is unjustifiable to approve the protocol or to conduct the experiment without this additional inquiry, which affects conditions 4 and 5 as well as 3.

Finally, consider this example: The principle of respect for autonomy and the principle of beneficence (which requires acts intended to prevent harm to others) sometimes come into contingent conflict when addressing situations that arise in governmental and professional responses to serious infectious-disease outbreaks, such as severe acquired respiratory syndrome (SARS). Persons exposed to SARS may put other persons at risk. The

government, under its public health responsibilities, and various health professionals have an obligation based on beneficence and justice to protect unexposed persons whenever possible. However, respect for autonomy often sets a prima facie barrier to infringements of liberty and privacy even in the context of public health concerns. To justify overriding respect for autonomy, one must show that mandatory quarantine of exposed individuals is necessary to prevent harm to others and has a reasonable prospect of preventing such harm. If it meets these conditions, mandatory quarantine still must pass the least-infringement test (condition 4), and public health officials should seek to minimize the negative effects of the quarantine, including the loss of income and the inability to care for dependent family members (condition 5). Finally, impartial application of the quarantine rules is essential for both fairness and public trust (condition 6).[42]

In our judgment, these six constraining conditions are morally demanding, at least in some circumstances. When conjoined with requirements of coherence presented in Chapter 10 (pp. 439–44), these conditions provide protections against purely intuitive, subjective, or biased balancing judgments. We could introduce further criteria or safeguards, such as "rights override nonrights" and "liberty principles override nonliberty principles," but these provisions are certain to fail in circumstances in which rights claims and liberty interests are relatively minor.

Moral Diversity and Moral Disagreement

Sometimes conscientious and reasonable moral agents understandably disagree over moral priorities in circumstances of a contingent conflict of norms. Morally conscientious persons may disagree, for example, about whether disclosure of a life-threatening condition to a fragile patient is appropriate, whether religious values about brain death have a place in secular biomedical ethics, whether mature teenagers should be permitted to refuse life-sustaining treatments, and other issues. Disagreement does not indicate moral ignorance or moral defect. We simply lack a single, entirely reliable way to resolve many disagreements, despite methods of specifying and balancing.

Moral disagreement can emerge because of (1) factual disagreements (e.g., about the level of suffering that an intervention will cause), (2) disagreements resulting from insufficient information or evidence, (3) disagreements about which norms are applicable or relevant in the circumstances, (4) disagreements about the relative weights or rankings of the relevant norms, (5) disagreements about appropriate forms of specification or balancing, (6) the presence of a genuine moral dilemma, (7) scope and moral status disagreements about who should be protected by a moral norm (e.g., whether embryos, fetuses, and sentient animals are protected; see Chapter 3), and (8) conceptual disagreements about a crucial moral concept such as whether removal of nutrition and hydration from a dying patient at a family's request constitutes *killing*.

Different parties may emphasize different principles or assign different weights to principles even when they agree on which principles and concepts are relevant. Disagreement may persist among morally committed persons who appropriately appreciate the basic demands that morality makes on them. If evidence is incomplete and different items of evidence are available to different parties, one individual or group may be justified in reaching a conclusion that another individual or group is justified in rejecting. Even if both parties have some incorrect beliefs, each party may have good reasons for holding those beliefs. We cannot hold persons to a higher practical standard than to make judgments conscientiously in light of the available norms and evidence.

When moral disagreements arise, a moral agent can—and usually should—defend his or her decision without disparaging or reproaching others who reach different decisions. Recognition of legitimate diversity—by contrast to moral violations that warrant criticism—is vital in the evaluation of the actions of others. One person's conscientious assessment of his or her obligations may differ from another's when they confront the same moral problem, and both evaluations may be appropriately grounded in the common morality. Similarly, what one institution or government determines it should do may differ from what another institution or government determines it should do. In such cases we can assess one position as morally preferable to another only if we can show that the position rests on a more coherent set of specifications and interpretations of the common morality.[43]

CONCLUSION

In this chapter we have presented what is sometimes called the *four-principles approach* to biomedical ethics, now commonly called *principlism*.[44] The four clusters of principles in our moral framework descend from the common morality, but when specifying and balancing these principles in later chapters we will also call on historical experience in formulating professional obligations and virtues in health care, public health, biomedical research, and health policy. Although various assumptions in traditional medical ethics, current medical and research codes, and other parts of contemporary bioethics need further reform, we are deeply indebted to their insights and commitments. Our goal in later chapters is to develop, specify, and balance the normative content of the four clusters of principles, and we will often seek to render our views consistent with professional traditions, practices, and codes.

Principlism is not merely a list of four abstract principles. It is a theory about how these principles are linked to and guide practice. In the nine chapters hereafter we show how principles and other moral norms are connected to an array of understandings, practices, and transactions in health care settings, research institutions, and public health policies.

NOTES

1. 1. See Albert Jonsen, *The Birth of Bioethics* (New York: Oxford University Press, 1998), pp. 3ff; Jonsen, *A Short History of Medical Ethics* (New York: Oxford University Press, 2000); John-Stewart Gordon, "Bioethics," in the *Internet Encyclopedia of Philosophy*, especially section 2, available at https://www.iep.utm.edu/bioethics/ (accessed March 23, 2018); and Edmund D. Pellegrino and David C. Thomasma, *The Virtues in Medical Practice* (New York: Oxford University Press, 1993), pp. 184–89.
2. 2. A comprehensive treatment of this history that ranges worldwide is Robert B. Baker and Laurence McCullough, eds., *The Cambridge World History of Medical Ethics* (Cambridge: Cambridge University Press, 2009).
3. 3. The language of "applied ethics" can be misleading insofar as it suggests one-way traffic from ethical theory and principles and rules to particular judgments about cases. In fact, particular case judgments interact dialectically with and may lead to modifications of theories, principles, and rules. See our discussion in Chapter 10, pp. 404–10.
4. 4. These distinctions should be used with caution. Metaethics frequently takes a turn toward the normative, and normative ethics often relies on metaethics. Just as no sharp distinction should be drawn between practical ethics and general normative ethics, no bright line should be drawn to distinguish normative ethics and metaethics.
5. 5. Although there is only one universal common morality, there is more than one theory of the common morality. For a diverse group of theories, see Alan Donagan, *The Theory of Morality* (Chicago: University of Chicago Press, 1977); Bernard Gert, *Common Morality: Deciding What to Do* (New York: Oxford University Press, 2007); Bernard Gert, Charles M. Culver, and K. Danner Clouser, *Bioethics: A Return to Fundamentals,* 2nd ed. (New York: Oxford University Press, 2006); W. D. Ross, *The Foundations of Ethics* (Oxford: Oxford University Press, 1939); and the special issue of the *Kennedy Institute of Ethics Journal* 13 (2003), especially the introductory article by Robert Veatch, pp. 189–92.

   For challenges to these theories and their place in bioethics, see John D. Arras, "The Hedgehog and the Borg: Common Morality in Bioethics," *Theoretical Medicine and Bioethics 30 (2009): 11–30; Arras, "A Common Morality for Hedgehogs: Bernard Gert's Method," in Arras, Methods in Bioethics: The Way We Reason Now*, ed. James F. Childress and Matthew Adams (New York: Oxford University Press, 2017), pp. 27–44; B. Bautz, "What Is the Common Morality, Really?" *Kennedy Institute of Ethics Journal* 26 (2016): 29–45; Carson Strong, "Is There No Common Morality?" *Medical Humanities Review* 11 (1997): 39–45; and Andrew Alexandra and Seumas Miller, "Ethical Theory, 'Common Morality,' and Professional Obligations," *Theoretical Medicine and Bioethics* 30 (2009): 69–80.

6. 6. See Martha Nussbaum's thesis that in Aristotle's philosophy, certain "non-relative virtues" are objective and universal. "Non-Relative Virtues: An Aristotelian Approach," in *Ethical Theory, Character, and Virtue,* ed. Peter French et al. (Notre Dame, IN: University of Notre Dame Press, 1988), pp. 32–53, especially pp. 33–4, 46–50. In a classic work in philosophical ethics, David Hume presents a theory of the virtues as objective and universal, though his theory is somewhat different from Aristotle's. See Hume's *An Enquiry concerning the Principles of Morals*, ed. Tom L. Beauchamp, in the series "Oxford Philosophical Texts Editions" (Oxford: Oxford University Press, 1998).
7. 7. For a broad and engaging account of common morality, see Rebecca Kukla, "Living with Pirates: Common Morality and Embodied Practice," *Cambridge Quarterly of Healthcare Ethics* 23 (2014): 75–85. See also Bernard Gert's insistence on the role of the *whole moral system* (not merely rules of obligation) and the perils of neglecting it, an often overlooked point with which we agree. See Gert's *Morality: Its Nature and Justification* (New York: Oxford University Press, 2005), pp. 3, 159–61, 246–47; and see also his "The Definition of Morality," in *The Stanford Encyclopedia of Philosophy*; revision of February 8, 2016, available at https://plato.stanford.edu/entries/morality-definition/ (accessed February 9, 2018).
8. 8. This mistaken interpretation of our theory is found in Leigh Turner, "Zones of Consensus and Zones of Conflict: Questioning the 'Common Morality' Presumption in Bioethics," *Kennedy Institute of Ethics Journal* 13 (2003): 193–218; and Turner, "An Anthropological Exploration of Contemporary Bioethics: The Varieties of Common Sense," *Journal of Medical Ethics* 24 (1998): 127–33.
9. 9. See David DeGrazia, "Common Morality, Coherence, and the Principles of Biomedical Ethics," *Kennedy Institute of Ethics Journal* 13 (2003): 219–30; Turner, "Zones of Consensus and Zones of Conflict"; Donald C. Ainslee, "Bioethics and the Problem of Pluralism," *Social Philosophy and Policy* 19 (2002): 1–28; Oliver Rauprich, "Common Morality: Comment on Beauchamp and Childress," *Theoretical Medicine and Bioethics* 29 (2008): 43–71; and Letícia Erig Osório de Azambuja and Volnei Garrafa, "The Common Morality Theory in the Work of Beauchamp and Childress," *Revista Bioética* 23 (2015), available at http://www.scielo.br/scielo.php?pid=S1983-80422015000300634&script=sci_arttext&tlng=en (accessed March 22, 2018). For a related, but distinguishable, criticism, see Anna E. Westra, Dick L. Willems, and Bert J. Smit, "Communicating with Muslim Parents: 'The Four Principles' Are not as Culturally Neutral as Suggested," *European Journal of Pediatrics* 168 (2009): 1383–87; this article is published together with a beautifully correct interpretation of our position by Voo Teck Chuan, "Editorial Comment: The Four Principles and Cultural Specification," *European Journal of Pediatrics* 168 (2009): 1389.
10. 10. Kukla reaches this conclusion in "Living with Pirates." See, in response, Tom L. Beauchamp, "On Common Morality as Embodied Practice: A Reply to Kukla," *Cambridge Quarterly of Healthcare Ethics* 23 (2014): 86–93; Carson Strong, "Kukla's Argument against Common Morality as a Set of Precepts: On Stranger Tides," *Cambridge Quarterly of Healthcare Ethics* 23 (2014): 93–99; and Kukla, "Response to Strong and Beauchamp—at World's End," *Cambridge Quarterly of Healthcare Ethics* 23 (2014): 99–102.
11. 11. See Richard B. Brandt, "Morality and Its Critics," in his *Morality, Utilitarianism, and Rights* (Cambridge: Cambridge University Press, 1992), chap. 5; and Gregory Mellema, "Moral Ideals and Virtue Ethics," *Journal of Ethics* 14 (2010): 173–80. See also our discussion of moral ideals and supererogation in Chapter 2, pp. 45–49.
12. 12. Talcott Parsons, *Essays in Sociological Theory,* rev. ed. (Glencoe, IL: Free Press, 1954), p. 372. See further Jan Nolin, *In Search of a New Theory of Professions* (Borås, Sweden: University of Borås, 2008).
13. 13. See the excellent introduction to this subject in Edmund D. Pellegrino, "Codes, Virtues, and Professionalism," in *Methods of Bioethics,* ed. Daniel Sulmasy and Jeremy Sugarman, 2nd ed. (Washington, DC: Georgetown University Press, 2010), pp. 91–108. For an overview of codes of medical ethics, see Robert Baker, "Medical Codes and Oaths," *Bioethics* [Formerly *Encyclopedia of Bioethics*], 4th ed., ed. Bruce Jennings (Farmington Hills, MI: Gale, Cengage Learning, Macmillan Reference USA, 2014), vol. 4, pp. 1935–46. For a history and assessment of the Code of Ethics for Nurses of the American Nurses Association, see Beth Epstein and Martha Turner, "The Nursing Code of Ethics: Its Value, Its History," *Online Journal of Issues in Nursing* 20, no. 2 (May 2015), available at http://ojin.nursingworld.org/MainMenuCategories/ANAMarketplace/ANAPeriodicals/OJIN/TableofContents/Vol-20-2015/No2-May-2015/The-Nursing-Code-of-Ethics-Its-Value-Its-History.html (accessed June 3, 2018).

14. 14. The American Medical Association Code of Ethics of 1847 was largely adapted from Thomas Percival's *Medical Ethics; or a Code of Institutes and Precepts, Adapted to the Professional Conduct of Physicians and Surgeons* (Manchester, UK: S. Russell, 1803). See Donald E. Konold, *A History of American Medical Ethics 1847–1912* (Madison, WI: State Historical Society of Wisconsin, 1962), chaps. 1–3; Chester Burns, "Reciprocity in the Development of Anglo-American Medical Ethics," in *Legacies in Medical Ethics,* ed. Burns (New York: Science History Publications, 1977); and American Medical Association, "History of the Code," available at https://www.ama-assn.org/sites/default/files/media-browser/public/ethics/ama-code-ethics-history.pdf (accessed March 23, 2018).
15. 15. For a related and rigorous critical analysis of Hippocratic and other medical codes, see Robert M. Veatch's influential views in his *Hippocratic, Religious, and Secular Medical Ethics: The Points of Conflict* (Washington, DC: Georgetown University Press, 2012).
16. 16. Cf. the conclusions reached about medicine in N. D. Berkman, M. K. Wynia, and L. R. Churchill, "Gaps, Conflicts, and Consensus in the Ethics Statements of Professional Associations, Medical Groups, and Health Plans," *Journal of Medical Ethics* 30 (2004): 395–401; Ryan M. Antiel, Farr A. Curlin, C. Christopher Hook, and Jon C. Tilburt, "The Impact of Medical School Oaths and Other Professional Codes of Ethics: Results of a National Physician Survey," *Archives of Internal Medicine* 171 (2011): 469–71; Robert D. Orr, Norman Pang, Edmund D. Pellegrino, and Mark Siegler, "Use of the Hippocratic Oath: A Review of Twentieth Century Practice and a Content Analysis of Oaths Administered in Medical Schools in the U.S. and Canada in 1993," *Journal of Clinical Ethics* 8 (1997): 377–88; and A. C. Kao and K. P. Parsi, "Content Analyses of Oaths Administered at U.S. Medical Schools in 2000," *Academic Medicine* 79 (2004): 882–87.
17. 17. Jay Katz, ed., *Experimentation with Human Beings* (New York: Russell Sage Foundation, 1972), pp. ix–x.
18. 18. For an examination of different models of public bioethics, see James F. Childress, "Reflections on the National Bioethics Advisory Commission and Models of Public Bioethics," *Goals and Practice of Public Bioethics: Reflections on National Bioethics Commissions*, special report, *Hastings Center Report* 47, no. 3 (2017): S20–S23, and several other essays in this special report. See also *Society's Choices: Social and Ethical Decision Making in Biomedicine*, ed. Ruth Ellen Bulger, Elizabeth Meyer Bobby, and Harvey V. Fineberg, for the Committee on the Social and Ethical Impacts of Developments in Biomedicine, Division of Health Sciences Policy, Institute of Medicine (Washington, DC: National Academies Press, 1995).
19. 19. See Allen Buchanan, "Philosophy and Public Policy: A Role for Social Moral Epistemology," *Journal of Applied Philosophy* 26 (2009): 276–90; Will Kymlicka, "Moral Philosophy and Public Policy: The Case of New Reproductive Technologies," in *Philosophical Perspectives on Bioethics,* ed. L. W. Sumner and Joseph Boyle (Toronto: University of Toronto Press, 1996); Dennis Thompson, "Philosophy and Policy," *Philosophy & Public Affairs* 14 (Spring 1985): 205–18; Andrew I. Cohen, *Philosophy, Ethics, and Public Policy* (London: Routledge, 2015); and a symposium on "The Role of Philosophers in the Public Policy Process: A View from the President's Commission," with essays by Alan Weisbard and Dan Brock, *Ethics* 97 (July 1987): 775–95.
20. 20. *Tarasoff v. Regents of the University of California,* 17 Cal. 3d 425, 551 P.2d 334, 131 Cal. Rptr. 14 (Cal. 1976).
21. 21. On the interactions of ethical and legal judgments (and the reasons for their interactions) on bioethical issues, see Stephen W. Smith, John Coggan, Clark Hobson, et al., eds., *Ethical Judgments: Re-Writing Medical Law* (Oxford: Hart, 2016).
22. 22. See John Lemmon, "Moral Dilemmas," *Philosophical Review* 71 (1962): 139–58; Daniel Statman, "Hard Cases and Moral Dilemmas," *Law and Philosophy* 15 (1996): 117–48; Terrance McConnell, "Moral Dilemmas," *Stanford Encyclopedia of Philosophy* (Fall 2014 edition), ed. Edward N. Zalta, available at https://plato.stanford.edu/archives/fall2014/entries/moral-dilemmas/ (accessed March 23, 2018); H. E. Mason, "Responsibilities and Principles: Reflections on the Sources of Moral Dilemmas," in *Moral Dilemmas and Moral Theory,* ed. H. E. Mason (New York: Oxford University Press, 1996).
23. 23. Christopher W. Gowans, ed., *Moral Dilemmas* (New York: Oxford University Press, 1987); Walter Sinnott-Armstrong, *Moral Dilemmas* (Oxford: Basil Blackwell, 1988); Edmund N. Santurri, *Perplexity in the Moral Life: Philosophical and Theological Considerations* (Charlottesville: University Press of Virginia, 1987). For an approach to dilemmas offered as an addition to our account in this chapter, see

Joseph P. DeMarco, "Principlism and Moral Dilemmas: A New Principle," *Journal of Medical Ethics* 31 (2005): 101–5.

24. 24. Some writers in biomedical ethics express reservations about the place of the particular principles we propose in this book. See Pierre Mallia, *The Nature of the Doctor–Patient Relationship*: *Health Care Principles through the Phenomenology of Relationships with Patients* (Springer Netherlands: Springer Briefs in Ethics, 2013), esp. chap. 2, "Critical Overview of Principlist Theories"; K. Danner Clouser and Bernard Gert, "A Critique of Principlism," *Journal of Medicine and Philosophy* 15 (April 1990): 219–36; Søren Holm, "Not Just Autonomy—The Principles of American Biomedical Ethics," *Journal of Medical Ethics* 21 (1994): 332–38; Peter Herissone-Kelly, "The Principlist Approach to Bioethics, and Its Stormy Journey Overseas," in *Scratching the Surface of Bioethics,* ed. Matti Häyry and Tuija Takala (Amsterdam: Rodopi, 2003), pp. 65–77; and numerous essays in *Principles of Health Care Ethics,* ed. Raanan Gillon and Ann Lloyd (London: Wiley, 1994); and *Principles of Health Care Ethics,* 2nd ed., ed. Richard E. Ashcroft et al. (Chichester, UK: Wiley, 2007).
25. 25. Thomas Percival, *Medical Ethics; or a Code of Institutes and Precepts, Adapted to the Professional Interests of Physicians and Surgeons* (Manchester: S. Russell, 1803 [and numerous later editions]). For commentary on this classic work and its influence, see Edmund D. Pellegrino, "Percival's Medical Ethics: The Moral Philosophy of an 18th-Century English Gentleman," *Archives of Internal Medicine* 146 (1986): 2265–69; Pellegrino, "Thomas Percival's Ethics: The Ethics Beneath the Etiquette" (Washington DC: Georgetown University, Kennedy Institute of Ethics, 1984), available at https://repository.library.georgetown.edu/bitstream/handle/10822/712018/Pellegrino_M269.pdf?sequence=1&isAllowed=n (accessed March 24, 2018); Robert B. Baker, Arthur L. Caplan, Linda L. Emanuel, and Stephen R. Latham, eds., *The American Medical Ethics Revolution: How the AMA's Code of Ethics Has Transformed Physicians' Relationships to Patients, Professionals, and Society* (Baltimore: Johns Hopkins University Press, 1999).
26. 26. Procedural rules might also be interpreted as grounded in substantive rules of equality. If so interpreted, the procedural rules could be said to have a justification in substantive rules.
27. 27. For a discussion of the distinction between *pro tanto* and *prima facie,* see Shelly Kagan, *The Limits of Morality* (Oxford: Clarendon Press, 1989), p. 17. Kagan prefers *pro tanto,* rather than prima facie, and notes that Ross used *prima facie* with effectively the same meaning, which some writers classify as a mistake on Ross's part. See further Andrew E. Reisner, "Prima Facie and Pro Tanto Oughts," *International Encyclopedia of Ethics* [online], first published February 1, 2013, available at https://onlinelibrary.wiley.com/doi/full/10.1002/9781444367072.wbiee406 (accessed March 24, 2018).
28. 28. W. D. Ross, *The Right and the Good* (Oxford: Clarendon Press, 1930), esp. pp. 19–36, 88. On important cautions about both the meaning and use of the related notion of "prima facie rights," see Joel Feinberg, *Rights, Justice, and the Bounds of Liberty* (Princeton, NJ: Princeton University Press, 1980), pp. 226–29, 232; and Judith Jarvis Thomson, *The Realm of Rights* (Cambridge, MA: Harvard University Press, 1990), pp. 118–29.
29. 29. Robert Nozick, "Moral Complications and Moral Structures," *Natural Law Forum* 13 (1968): 1–50, available at https://scholarship.law.nd.edu/cgi/viewcontent.cgi?article=1136...naturallaw_forum (accessed March 26, 2018); James J. Brummer, "Ross and the Ambiguity of Prima Facie Duty," *History of Philosophy Quarterly* 19 (2002): 401–22. See also Thomas E. Hill, Jr., "Moral Dilemmas, Gaps, and Residues: A Kantian Perspective"; Walter Sinnott-Armstrong, "Moral Dilemmas and Rights"; and Terrance C. McConnell, "Moral Residue and Dilemmas"—all in *Moral Dilemmas and Moral Theory,* ed. Mason.
30. 30. For a similar view, see Ross, *The Right and the Good,* p. 28.
31. 31. Henry S. Richardson, "Specifying Norms as a Way to Resolve Concrete Ethical Problems," *Philosophy & Public Affairs* 19 (Fall 1990): 279–310; and Richardson, "Specifying, Balancing, and Interpreting Bioethical Principles," *Journal of Medicine and Philosophy* 25 (2000): 285–307, also in *Belmont Revisited: Ethical Principles for Research with Human Subjects,* ed. James F. Childress, Eric M. Meslin, and Harold T. Shapiro (Washington, DC: Georgetown University Press, 2005), pp. 205–27. See also David DeGrazia, "Moving Forward in Bioethical Theory: Theories, Cases, and Specified Principlism," *Journal of Medicine and Philosophy* 17 (1992): 511–39.
32. 32. Richardson, "Specifying, Balancing, and Interpreting Bioethical Principles," p. 289.

33. 33. For an excellent critical examination and case study of how the four-principles framework and approach can and should be used as a practical instrument, see John-Stewart Gordon, Oliver Rauprich, and Jochen Vollmann, "Applying the Four-Principle Approach," *Bioethics* 25 (2011): 293–300, with a reply by Tom Beauchamp, "Making Principlism Practical: A Commentary on Gordon, Rauprich, and Vollmann," *Bioethics* 25 (2011): 301–3.
34. 34. American Academy of Psychiatry and the Law, "Ethical Guidelines for the Practice of Forensic Psychiatry," as revised and adopted May 2005, section III: "The informed consent of the person undergoing the forensic evaluation should be obtained when necessary and feasible. If the evaluee is not competent to give consent, the evaluator should follow the appropriate laws of the jurisdiction. … [P]sychiatrists should inform the evaluee that if the evaluee refuses to participate in the evaluation, this fact may be included in any report or testimony. If the evaluee does not appear capable of understanding the information provided regarding the evaluation, this impression should also be included in any report and, when feasible, in testimony." Available at http://www.aapl.org/ethics.htm (accessed February 19, 2018).
35. 35. Dennis H. Novack et al., "Physicians' Attitudes toward Using Deception to Resolve Difficult Ethical Problems," *Journal of the American Medical Association* 261 (May 26, 1989): 2980–85. We return to these problems in Chapter 8 (pp. 327–37).
36. 36. Richardson, "Specifying Norms," p. 294. The word "always" in this formulation should be understood to mean "in principle always." Specification may, in some cases, reach a final form.
37. 37. Other prohibitions, such as rules against murder and rape, may be absolute only because of the meaning of their terms. For example, to say "murder is categorically wrong" may be only to say "unjustified killing is unjustified."
38. 38. Timothy Quill and Penelope Townsend, "Bad News: Delivery, Dialogue, and Dilemmas," *Archives of Internal Medicine* 151 (March 1991): 463–68.
39. 39. See Alisa Carse, "Impartial Principle and Moral Context: Securing a Place for the Particular in Ethical Theory," *Journal of Medicine and Philosophy* 23 (1998): 153–69. For a defense of balancing as the best method in such situations, see Joseph P. DeMarco and Paul J. Ford, "Balancing in Ethical Deliberations: Superior to Specification and Casuistry," *Journal of Medicine and Philosophy* 31 (2006): 483–97, esp. 491–93.
40. 40. See similar reflections in Lawrence Blum, *Moral Perception and Particularity* (New York: Cambridge, 1994), p. 204.
41. 41. To the extent these six conditions incorporate moral norms, the norms are prima facie, not absolute. Condition 3 is redundant if it cannot be violated when all of the other conditions are satisfied; but it is best to be clear on this point, even if redundant.
42. 42. See James F. Childress and Ruth Gaare Bernheim, "Public Health Ethics: Public Justification and Public Trust," *Bundesgundheitsblat: Gusundheitsforschung, Gesundheitsschutz* 51, no. 2 (February 2008): 158–63; and Ruth Gaare Bernheim, James F. Childress, Richard J. Bonnie, and Alan L. Melnick, *Essentials of Public Health Ethics: Foundations, Tools, and Interventions* (Boston: Jones and Bartlett, 2014), esp. chaps. 1, 2, and 8.
43. 43. For a criticism of our conclusion in this paragraph, see Marvin J. H. Lee, "The Problem of 'Thick in Status, Thin in Content,' in Beauchamp and Childress's Principlism," *Journal of Medical Ethics* 36 (2010): 525–28. See further Angus Dawson and E. Garrard, "In Defence of Moral Imperialism: Four Equal and Universal Prima Facie Principles," *Journal of Medical Ethics* 32 (2006): 200–204; Walter Sinnott-Armstrong, *Moral Dilemmas,* pp. 216–27; and D. D. Raphael, *Moral Philosophy* (Oxford: Oxford University Press, 1981), pp. 64–65.
44. 44. See Bernard Gert, Charles M. Culver, and K. Danner Clouser, *Bioethics: A Return to Fundamentals,* 2nd ed., chap. 4; Clouser and Gert, "A Critique of Principlism," pp. 219–36; Carson Strong, "Specified Principlism," *Journal of Medicine and Philosophy* 25 (2000): 285–307; John H. Evans, "A Sociological Account of the Growth of Principlism," *Hastings Center Report* 30 (September–October 2000): 31–38; Evans, *Playing God: Human Genetic Engineering and the Rationalization of Public Bioethical Debate* (Chicago: University of Chicago Press, 2002); and Evans, *The History and Future of Bioethics: A Sociological View* (New York: Oxford University Press, 2011). For a critical analysis of Evans's arguments, particularly in *Playing God*, see James F. Childress, "Comments," *Journal of the Society of Christian Ethics* 24, no. 1 (2004): 195–204.

2

Moral Character

Chapter 1 concentrated on moral norms in the form of principles, rules, obligations, and rights. This chapter focuses on moral character, especially moral virtues, moral ideals, and moral excellence. These categories complement those in the previous chapter. The moral norms discussed in Chapter 1 chiefly govern right and wrong *action*. By contrast, character ethics and virtue ethics concentrate on the *agent* who performs actions and the virtues that make agents morally worthy persons.[1]

The goals and structure of medicine, health care, public health, and research call for a deep appreciation of moral virtues. What often matters most in health care interactions and in the moral life generally is not adherence to moral rules but having a reliable character, good moral sense, and appropriate emotional responsiveness. Even carefully specified principles and rules do not convey what occurs when parents lovingly play with and nurture their children or when physicians and nurses exhibit compassion, patience, and responsiveness in their encounters with patients and families. The feelings and concerns for others that motivate us to take actions often cannot be reduced to a sense of obligation to follow rules. Morality would be a cold and uninspiring practice without appropriate sympathy, emotional responsiveness, excellence of character, and heartfelt ideals that reach beyond principles and rules.

Some philosophers have questioned the place of virtues in moral theory. They see virtues as less central than action-guiding norms and as difficult to unify in a systematic theory, in part because there are many independent virtues to be considered. Utilitarian Jeremy Bentham famously complained that there is "no marshaling" the virtues and vices because "they are susceptible of no arrangement; they are a disorderly body, whose members are frequently in hostility with one another. … Most of them are characterized by that vagueness which is a convenient instrument for the poetical, but dangerous or useless to the practical moralist."[2]

Although principles and virtues are different and learned in different ways, virtues are no less important in the moral life, and in some contexts are probably more important. In Chapter 9, we examine virtue ethics as a type of moral theory and address challenges and criticisms such as Bentham's. In the first few sections of the present chapter, we analyze the concept of virtue; examine virtues in professional roles; treat the moral virtues of care, caregiving, and caring in health care; and explicate five other focal virtues in both health care and research.

The Concept of Moral Virtue

A *virtue* is a dispositional trait of character that is socially valuable and reliably present in a person, and a *moral virtue* is a dispositional trait of character that is morally valuable and reliably present. If cultures or social groups approve a trait and regard it as moral, their approval is not sufficient to qualify the trait as a moral virtue. Moral virtue is more than a personal, dispositional trait that is socially approved in a particular group or culture.[3] This approach to the moral virtues accords with our conclusion in Chapter 1 that the common morality excludes provisions found only in so-called cultural moralities and individual moralities. The moral virtues, like moral principles, are part of the common morality.

Some define the term *moral virtue* as a disposition to act or a habit of acting in accordance with, and with the aim of following, moral principles, obligations, or ideals.[4] For example, they understand the moral virtue of nonmalevolence as the trait of abstaining from causing harm to others when it would be wrong to cause harm. However, this definition unjustifiably views virtues as merely derivative from and dependent on principles and fails to capture the importance of moral motives. We care morally about people's motives, and we care especially about their characteristic motives and dispositions, that is, the motivational structures embedded in their character. Persons who are motivated through impartial sympathy and personal affection, for example, are

likely to meet our moral approval, whereas persons who act similarly, but are motivated merely by personal ambition, do not.

Consider a person who discharges moral obligations only because they are moral requirements while intensely disliking being obligated to place the interests of others above his or her personal interests and projects. This person does not feel friendly toward or cherish others and respects their wishes only because moral obligation requires it. If this person's motive is deficient, a critical moral ingredient is missing even though he or she consistently performs morally right actions and has a disposition to perform right actions. When a person characteristically lacks an appropriate motivational structure, a necessary condition of virtuous character is absent. The act may be right and the actor blameless, but neither the act nor the actor is *virtuous*. People may be disposed to do what is right, intend to do it, and do it, while simultaneously yearning to avoid doing it. Persons who characteristically perform morally right actions from such a motivational structure are not morally virtuous even if they invariably perform the morally right action.

Such a person has a morally deficient character, and he or she performs morally right actions for reasons or feelings disconnected from moral motivation. A philanthropist's gift of a new wing of a hospital will be recognized by hospital officials and by the general public as a generous gift, but if the philanthropist is motivated only by a felt need for public praise and only makes the gift to gain such praise, there is a discordance between those feelings and the performance of the praised action. Feelings, intentions, and motives are morally important in a virtue theory in a way that may be lost or obscured in an obligation-based theory.[5]

VIRTUES IN PROFESSIONAL ROLES

Persons differ in their sets of character traits. Most individuals have some virtues and some vices while lacking other virtues and vices. However, all persons with normal moral capacities can cultivate the character traits centrally important to morality such as honesty, fairness, fidelity, truthfulness, and benevolence. In professional life in health care and research, the traits that warrant encouragement and admiration often derive from role responsibilities. Some virtues are essential for enacting these professional roles, and certain vices are intolerable in professional life. Accordingly, we turn now to virtues that are critically important in professional and institutional roles and practices in biomedical fields.

Virtues in Roles and Practices

Professional roles are grounded in institutional expectations and governed by established standards of professional practice. Roles internalize conventions, customs, and procedures of teaching, nursing, doctoring, and the like. Professional practice has traditions that require professionals to cultivate certain virtues. Standards of virtue incorporate criteria of professional merit, and possession of these virtues disposes persons to act in accordance with the objectives of the practices.

In the practice of medicine, several goods internal to the profession are appropriately associated with being a good physician. These goods include specific moral and nonmoral skills in the care of patients, the application of specific forms of knowledge, and the teaching of health behaviors. They are achievable only if one lives up to the standards of the good physician, standards that in part define the practice. A practice is not merely a set of technical skills. Practices should be understood in terms of the respect that practitioners have for the goods internal to the practices. Although these practices sometimes need to be revised, the historical development of a body of standards has established many practices now found at the heart of medicine, nursing, and public health.[6]

Roles, practices, and virtues in medicine, nursing, and other health care and research professions reflect social expectations as well as standards and ideals internal to these professions.[7] The virtues we highlight in this chapter are care—a fundamental virtue for health care relationships—along with five focal virtues found in all health care professions: compassion, discernment, trustworthiness, integrity, and conscientiousness, all of which

support and promote caring and caregiving. Elsewhere in this chapter and in later chapters, we discuss other virtues, including respectfulness, nonmalevolence, benevolence, justice, truthfulness, and fidelity.

To illustrate the difference between standards of moral character in a profession and standards of technical performance in a profession, we begin with an instructive study of surgical error. Charles L. Bosk's influential *Forgive and Remember: Managing Medical Failure* presents an ethnographic study of the way two surgical services handle medical failure, especially failures by surgical residents in "Pacific Hospital" (a name substituted for the hospitals actually studied).[8] Bosk found that both surgical services distinguish, at least implicitly, between several different forms of error or mistake. The first form is *technical*: A professional discharges role responsibilities conscientiously, but his or her technical training or information still falls short of what the task requires. Every surgeon will occasionally make this sort of mistake. A second form of error is *judgmental*: A conscientious professional develops and follows an incorrect strategy. These errors are also to be expected. Attending surgeons forgive momentary technical and judgmental errors but remember them in case a pattern develops indicating that a surgical resident lacks the technical and judgmental skills to be a competent surgeon. A third form of error is *normative*: A physician violates a norm of conduct or fails to possess a moral skill, particularly by failing to discharge moral obligations conscientiously or by failing to acquire and exercise critical moral virtues such as conscientiousness. Bosk concludes that surgeons regard technical and judgmental errors as less important than moral errors, because every conscientious person can be expected to make "honest errors" or "good faith errors," whereas moral errors such as failures of conscientiousness are considered profoundly serious when a pattern indicates a defect of character.

Bosk's study indicates that persons of high moral character acquire a reservoir of goodwill in assessments of either the praiseworthiness or the blameworthiness of their actions. If a conscientious surgeon and another surgeon who is not adequately conscientious make the same technical or judgmental errors, the conscientious surgeon will not be subjected to moral blame to the same degree as the other surgeon.

Virtues in Different Professional Models

Professional virtues were historically integrated with professional obligations and ideals in codes of health care ethics. Insisting that the medical profession's "prime objective" is to render service to humanity, an American Medical Association (AMA) code in effect from 1957 to 1980 urged the physician to be "upright" and "pure in character and … diligent and conscientious in caring for the sick." It endorsed the virtues that Hippocrates commended: modesty, sobriety, patience, promptness, and piety. However, in contrast to its first code of 1847, the AMA over the years has increasingly de-emphasized virtues in its codes. The 1980 version for the first time eliminated all trace of the virtues except for the admonition to expose "those physicians deficient in character or competence." This pattern of de-emphasis regrettably still continues.

Thomas Percival's 1803 book, *Medical Ethics*, is a classic example of an attempt to establish the proper set of virtues in medicine. Starting from the assumption that the patient's best medical interest is the proper goal of medicine, Percival reached conclusions about the good physician's traits of character, which were primarily tied to responsibility for the patient's medical welfare.[9] This model of medical ethics supported medical paternalism with effectively no attention paid to respect for patients' autonomous choices.

In traditional nursing, where the nurse was often viewed as the "handmaiden" of the physician, the nurse was counseled to cultivate the passive virtues of obedience and submission. In contemporary models in nursing, by contrast, active virtues have become more prominent. For example, the nurse's role is now often regarded as one of advocacy for patients.[10] Prominent virtues include respectfulness, considerateness, justice, persistence, and courage.[11] Attention to patients' rights and preservation of the nurse's integrity also have become increasingly prominent in some contemporary models.

The conditions under which ordinarily praiseworthy virtues become morally unworthy present thorny ethical issues. Virtues such as loyalty, courage, generosity, kindness, respectfulness, and benevolence at times lead persons to act inappropriately and unacceptably. For instance, the physician or nurse who acts kindly and loyally

by not reporting the incompetence of a fellow physician or nurse acts unethically. This failure to report misconduct does not suggest that loyalty and kindness are not virtues. It indicates only that the virtues need to be accompanied by an understanding of what is right and good and of what deserves loyalty, kindness, generosity, and the like.

THE CENTRAL VIRTUE OF CARING

As the language of *health care*, *medical care*, and *nursing care* suggests, the virtue of care, or caring, is prominent in professional ethics. We treat this virtue as fundamental in relationships, practices, and actions in health care. In explicating this family of virtues we draw on what has been called the *ethics of care*, which we interpret as a form of virtue ethics.[12] The ethics of care emphasizes traits valued in intimate personal relationships such as sympathy, compassion, fidelity, and love. *Caring* refers to care for, emotional commitment to, and willingness to act on behalf of persons with whom one has a significant relationship. *Caring for* is expressed in actions of "caregiving," "taking care of," and "due care." The nurse's or physician's trustworthiness and quality of care and sensitivity in the face of patients' problems, needs, and vulnerabilities are integral to their professional moral lives.

The ethics of care emphasizes what physicians and nurses do—for example, whether they break or maintain confidentiality—and how they perform those actions, which motives and feelings underlie them, and whether their actions promote or thwart positive relationships.

The Origins of the Ethics of Care

The ethics of care, understood as a form of philosophical ethics, originated and continues to flourish in feminist writings. The earliest works emphasized how women display an ethic of care, by contrast to men, who predominantly exhibit an ethic of rights and obligations. Psychologist Carol Gilligan advanced the influential hypothesis that "women speak in a different voice"—a voice that traditional ethical theory failed to appreciate. She discovered "the voice of care" through empirical research involving interviews with girls and women. This voice, she maintained, stresses empathic association with others, not based on "the primacy and universality of individual rights, but rather on … a very strong sense of being responsible."[13]

Gilligan identified two modes of moral thinking: an ethic of care and an ethic of rights and justice. She did not claim that these two modes of thinking strictly correlate with gender or that all women or all men speak in the same moral voice.[14] She maintained only that men tend to embrace an ethic of rights and justice that uses quasi-legal terminology and impartial principles, accompanied by dispassionate balancing and conflict resolution, whereas women tend to affirm an ethic of care that centers on responsiveness in an interconnected network of needs, care, and prevention of harm.[15]

Criticisms of Traditional Theories by Proponents of an Ethics of Care

Proponents of the care perspective often criticize traditional ethical theories that tend to de-emphasize virtues of caring. Two criticisms merit consideration here.[16]

Challenging impartiality. Some proponents of the care perspective argue that theories of obligation unduly telescope morality by overemphasizing detached fairness. This orientation is suitable for some moral relationships, especially those in which persons interact as equals in a public context of impersonal justice and institutional constraints, but moral detachment also may reflect a lack of caring responsiveness. In the extreme case, detachment becomes uncaring indifference. Lost in the *detachment* of impartiality is an *attachment* to what we care about most and is closest to us—for example, our loyalty to family, friends, and groups. Here partiality toward others is morally permissible and is an expected form of interaction. This kind of partiality is a feature of the human condition without which we might impair or sever our most important relationships.[17]

Proponents of a care ethics do not recommend complete abandonment of principles if principles are understood to allow room for discretionary and contextual judgment. However, some defenders of the ethics of care find principles largely irrelevant, ineffectual, or unduly constrictive in the moral life. A defender of principles could hold that principles of care, compassion, and kindness tutor our responses in caring, compassionate, and kind ways. But this attempt to rescue principles seems rather empty. Moral experience confirms that we often do rely on our emotions, capacity for sympathy, sense of friendship, and sensitivity to find appropriate moral responses. We could produce rough generalizations about how caring clinicians should respond to patients, but such generalizations cannot provide adequate guidance for all interactions. Each situation calls for responses beyond following rules, and actions that are caring in one context may be offensive or even harmful in another.

Relationships and emotion. The ethics of care places special emphasis on mutual interdependence and emotional responsiveness. Many human relationships in health care and research involve persons who are vulnerable, dependent, ill, and frail. Feeling for and being immersed in the other person are vital aspects of a moral relationship with them.[18] A person seems morally deficient if he or she acts according to norms of obligation without appropriately aligned feelings, such as concern and sympathy for a patient who is suffering. Good health care often involves insight into the needs of patients and considerate attentiveness to their circumstances.[19]

In the history of human experimentation, those who first recognized that some subjects of research were brutalized, subjected to misery, or placed at unjustifiable risk were persons able to feel sympathy, compassion, disgust, and outrage about the situation of these research subjects. They exhibited perception of and sensitivity to the feelings of subjects where others lacked comparable perceptions, sensitivities, and responses. This emotional sensitivity does not reduce moral response to emotional response. Caring has a cognitive dimension and requires a range of moral skills that involve insight into and understanding of another's circumstances, needs, and feelings.

One proponent of the ethics of care argues that action is sometimes appropriately principle-guided, but not necessarily always governed by or derived from principles.[20] This statement moves in the right direction for construction of a comprehensive moral framework. We need not reject principles of obligation in favor of virtues of caring, but moral judgment involves moral skills beyond those of specifying and balancing general principles. An ethic that emphasizes the virtues of caring well serves health care because it is close to the relationships and processes of decision making found in clinical contexts, and it provides insights into basic commitments of caring and caretaking. It also liberates health professionals from the narrow conceptions of role responsibilities that have been delineated in some professional codes of ethics.

FIVE FOCAL VIRTUES

We now turn to five focal virtues for health professionals: compassion, discernment, trustworthiness, integrity, and conscientiousness. These virtues are important for the development and expression of caring, which we have presented as a fundamental orienting virtue in health care. These five additional virtues provide a moral compass of character for health professionals that builds on centuries of thought about health care ethics.[21]

Compassion

Compassion, says Edmund Pellegrino, is a "prelude to caring."[22] The virtue of compassion combines an attitude of active regard for another's welfare together with sympathy, tenderness, and discomfort at another's misfortune or suffering.[23] Compassion presupposes sympathy, has affinities with mercy, and is expressed in acts of beneficence that attempt to alleviate the misfortune or suffering of another person.

Nurses and physicians must understand the feelings and experiences of patients to respond appropriately to them and their illnesses and injuries—hence the importance of empathy, which involves sensing or even reconstructing another person's mental experience, whether that experience is negative or positive.[24] As

important as empathy is for compassion and other virtues, the two are different, and empathy does not always lead to compassion. Some literature on professionalism in medicine and health care now focuses on empathy rather than compassion, but this literature risks making the mistake of viewing empathy alone as sufficient for humanizing medicine and health care while overlooking its potential dangers.[25]

Compassion generally focuses on others' pain, suffering, disability, or misery—the typical occasions for compassionate response in health care. Using the language of *sympathy*, eighteenth-century philosopher David Hume pointed to a typical circumstance of compassion in surgery and explained how such feelings arise:

> Were I present at any of the more terrible operations of surgery, 'tis certain, that even before it begun, the preparation of the instruments, the laying of the bandages in order, the heating of the irons, with all the signs of anxiety and concern in the patient and assistants, wou'd have a great effect upon my mind, and excite the strongest sentiments of pity and terror. No passion of another discovers itself immediately to the mind. We are only sensible of its causes or effects. From *these* we infer the passion: And consequently *these* give rise to our sympathy.[26]

Physicians and nurses who express little or no compassion in their behavior may fail to provide what patients need most. The physician, nurse, or social worker altogether lacking in the appropriate display of compassion has a moral weakness. However, compassion also can cloud judgment and preclude rational and effective responses. In one reported case, a long-alienated son wanted to continue a futile and painful treatment for his near-comatose father in an intensive care unit (ICU) to have time to "make his peace" with his father. Although the son understood that his alienated father had no cognitive capacity, the son wanted to work through his sense of regret and say a proper good-bye. Some hospital staff argued that the patient's grim prognosis and pain, combined with the needs of others waiting to receive care in the ICU, justified stopping the treatment, as had been requested by the patient's close cousin and informal guardian. Another group in the unit regarded continued treatment as an appropriate act of compassion toward the son, who they thought should have time to express his farewells and regrets to make himself feel better about his father's death. The first group, by contrast, viewed this expression of compassion as misplaced because of the patient's prolonged agony and dying. In effect, those in the first group believed that the second group's compassion prevented clear thinking about primary obligations to this patient.[27]

Numerous writers in the history of ethical theory have proposed a cautious approach to compassion. They argue that a passionate, or even a compassionate, engagement with others can blind reason and prevent impartial reflection. Health care professionals understand and appreciate this phenomenon. Constant contact with suffering can overwhelm and even paralyze a compassionate physician or nurse. Impartial judgment sometimes gives way to impassioned decisions, and emotional burnout can arise. To counteract this problem, medical education and nursing education are well designed when they inculcate detachment alongside compassion. The language of *detached concern* and *compassionate detachment* came to the fore in this context.

Discernment

The virtue of discernment brings sensitive insight, astute judgment, and understanding to bear on action. Discernment involves the ability to make fitting judgments and reach decisions without being unduly influenced by extraneous considerations, fears, personal attachments, and the like. Some writers closely associate discernment with practical wisdom, or *phronesis*, to use Aristotle's widely used term. A person of practical wisdom knows which ends to choose, knows how to realize them in particular circumstances, and carefully selects from among the range of possible actions, while keeping emotions within proper bounds. In Aristotle's model, the practically wise person understands how to act with the right intensity of feeling, in just the right way, at just the right time, with a proper balance of reason and desire.[28]

A discerning person is disposed to understand and perceive what circumstances demand in the way of human responsiveness. For example, a discerning physician will see when a despairing patient needs comfort rather than privacy, and vice versa. If comfort is the right choice, the discerning physician will find the right type and

level of consolation to be helpful rather than intrusive. If a rule guides action in a particular case, seeing *how* to best follow the rule involves a form of discernment that is independent of seeing *that* the rule applies.

Accordingly, the virtue of discernment involves understanding both that and how principles and rules apply. Acts of respect for autonomy and beneficence therefore will vary in health care contexts, and the ways in which clinicians discerningly implement these principles in the care of patients will be as different as the many ways in which devoted parents care for their children.

Trustworthiness

Virtues, Annette Baier maintains, "are personal traits that contribute to a good climate of trust between people, when trust is taken to be acceptance of being, to some degree and in some respects, in another's power."[29] Trust is a confident belief in and reliance on the moral character and competence of another person, often a person with whom one has an intimate or established relationship. Trust entails a confidence that another will reliably act with the right motives and feelings and in accordance with appropriate moral norms.[30] To be *trustworthy* is to warrant another's confidence in one's character and conduct.

Traditional ethical theories rarely mention either trust or trustworthiness. However, Aristotle took note of one important aspect of trust and trustworthiness. He maintained that when relationships are voluntary and among intimates, by contrast to legal relationships among strangers, it is appropriate for the law to forbid lawsuits for harms that occur. Aristotle reasoned that intimate relationships involving "dealings with one another as good and trustworthy" hold persons together more than "bonds of justice" do.[31]

Nothing is more valuable in health care organizations and contexts than the maintenance of a culture of trust. Trust and trustworthiness are essential when patients are vulnerable and place their hope and their confidence in health care professionals. A true climate of trust is endangered in contemporary health care institutions, as evidenced by the number of medical malpractice suits and adversarial relations between health care professionals and the public. Overt distrust has been engendered by mechanisms of managed care, because of the incentives some health care organizations create for physicians to limit the amount and kinds of care they provide to patients. Appeals have increased for ombudsmen, patient advocates, legally binding "directives" to physicians, and the like. Among the contributing causes of the erosion of a climate of trust are the loss of intimate contact between physicians and patients, the increased use of specialists, the lack of adequate access to adequate health care insurance, and the growth of large, impersonal, and bureaucratic medical institutions.[32]

Integrity

Some writers in bioethics hold that the primary virtue in health care is integrity.[33] People often justify their actions or refusals to act on grounds that they would otherwise compromise or sacrifice their integrity. Later in this chapter we discuss appeals to integrity as invocations of *conscience*, but we confine attention at present to the virtue of integrity.

The central place of integrity in the moral life is beyond dispute, but what the term means is less clear. In its most general sense, "moral integrity" means soundness, reliability, wholeness, and integration of moral character. In a more restricted sense, the term refers to objectivity, impartiality, and fidelity in adherence to moral norms. Accordingly, the virtue of integrity represents two aspects of a person's character. The first is a coherent integration of aspects of the self—emotions, aspirations, knowledge, and the like—so that each complements and does not frustrate the others. The second is the character trait of being faithful to moral values and standing up in their defense when necessary. A person can lack moral integrity in several respects—for example, through hypocrisy, insincerity, bad faith, and self-deception. These vices represent breaks in the connections among a person's moral convictions, emotions, and actions. The most common deficiency is probably a lack of sincerely and firmly held moral convictions, but no less important is the failure to act consistently on the moral beliefs that one does hold.

Problems in maintaining integrity may also arise from a conflict of moral norms, or from moral demands that require persons to halt or abandon personal goals and projects. Persons may experience a sense of loss of their autonomy and feel violated by the demand to sacrifice their personal commitments and objectives.[34] For example, if a nurse is the only person in her family who can properly manage her mother's health, health care, prescription medications, nursing home arrangements, explanations to relatives, and negotiations with physicians, little time may be left for her personal projects and commitments. Such situations can deprive persons of the liberty to structure and integrate their lives as they choose. If a person has structured his or her life around personal goals that are ripped away by the needs and agendas of others, a loss of personal integrity occurs.

Problems of professional integrity often center on wrongful conduct in professional life. When breaches of professional integrity involve violations of professional standards, they are viewed as violations of the rules of professional associations, codes of medical ethics, or medical traditions,[35] but this vision of integrity needs to be broadened. Breaches of professional integrity also occur when a physician prescribes a drug that is no longer recommended for the outcome needed, enters into a sexual relationship with a patient, or follows a living will that calls for a medically inappropriate intervention.

Sometimes conflicts arise between a person's sense of moral integrity and what is required for professional integrity. Consider medical practitioners who, because of their religious commitments to the sanctity of life, find it difficult to participate in decisions not to do everything possible to prolong life. To them, participating in removing ventilators and intravenous fluids from patients, even from patients with a clear advance directive, violates their moral integrity. Their commitments may create morally troublesome situations in which they must either compromise their fundamental commitments or withdraw from the care of the patient. Yet compromise seems what a person, or an organization, of integrity cannot do, because it involves the sacrifice of deep moral commitments.[36]

Health care facilities cannot entirely eliminate these and similar problems of staff disagreement and conflicting commitments, but persons with the virtues of patience, humility, and tolerance can help reduce the problems. Situations that compromise integrity can be ameliorated if participants anticipate the problem before it arises and recognize the limits and fallibility of their personal moral views. Participants in a dispute may also have recourse to consultative institutional processes, such as hospital ethics committees. However, it would be ill-advised to recommend that a person of integrity can and should always negotiate and compromise his or her values in an intrainstitutional confrontation. There is something ennobling and admirable about the person or organization that refuses to compromise beyond a certain carefully considered moral threshold. To compromise below the threshold of integrity is simply to lose it.

Conscientiousness

The subject of integrity and compromise leads directly to a discussion of the virtue of conscientiousness and accounts of conscience. An individual acts conscientiously if he or she is motivated to do what is right because it is right, has worked with due diligence to determine what is right, intends to do what is right, and exerts appropriate effort to do so. Conscientiousness is the character trait of acting in this way.

Conscience and conscientiousness. *Conscience* has often been viewed as a mental faculty of, and authority for, moral decision making.[37] Slogans such as "Let your conscience be your guide" suggest that conscience is the final authority in moral justification. However, such a view fails to capture the nature of either conscience or conscientiousness, as the following case presented by Bernard Williams helps us see: Having recently completed his PhD in chemistry, George has not been able to find a job. His family has suffered from his failure. They are short of money, his wife has had to take additional work, and their small children have been subjected to considerable strain, uncertainty, and instability. An established chemist can get George a position in a laboratory that pursues research on chemical and biological weapons. Despite his perilous financial and familial circumstances, George concludes that he cannot accept this position because of his conscientious opposition to chemical and biological warfare. The senior chemist notes that the research will continue no matter what George

decides. Furthermore, if George does not take this position, it will be offered to another young man who would vigorously pursue the research. Indeed, the senior chemist confides, his concern about the other candidate's nationalistic fervor and uncritical zeal for research in chemical and biological warfare motivated him to recommend George for the job. George's wife is puzzled and hurt by George's reaction. She sees nothing wrong with the research. She is profoundly concerned about their children's problems and the instability of their family. Nonetheless, George forgoes this opportunity both to help his family and to prevent a destructive fanatic from obtaining the position. He says his conscience stands in the way.[38]

Conscience, as this example suggests, is neither a special moral faculty nor a self-justifying moral authority. It is a form of self-reflection about whether one's acts are obligatory or prohibited, right or wrong, good or bad, virtuous or vicious. It involves an internal sanction that comes into play through critical reflection. When individuals recognize their acts as violations of an appropriate standard, this sanction often appears as a bad conscience in the form of feelings of remorse, guilt, shame, disunity, or disharmony. A conscience that sanctions conduct in this way does not signify bad moral character. To the contrary, this experience of conscience is most likely to occur in persons of strong moral character and may even be a necessary condition of morally good character.[39] Kidney donors have been known to say, "I had to do it. I couldn't have backed out, not that I had the feeling of being trapped, because the doctors offered to get me out. I just had to do it."[40] Such judgments derive from ethical standards that are sufficiently powerful that violating them would diminish integrity and result in guilt or shame.[41]

When people claim that their actions are conscientious, they sometimes feel compelled by conscience to resist others' authoritative demands. Instructive examples are found in military physicians who believe they must answer first to their consciences and cannot plead "superior orders" when commanded by a superior officer to commit what they believe to be a moral wrong. Agents sometimes act out of character in order to perform what they judge to be the morally appropriate action. For example, a normally cooperative and agreeable physician may indignantly, but justifiably, protest an insurance company's decision not to cover the costs of a patient's treatment. Such moral indignation and outrage can be appropriate and admirable.

Conscientious refusals. Conscientious objections and refusals by physicians, nurses, pharmacists, and other health care professionals raise difficult issues for public policy, professional organizations, and health care institutions. Examples are found in a physician's refusal to honor a patient's legally valid advance directive to withdraw artificial nutrition and hydration, a nurse's refusal to participate in an abortion or sterilization procedure, and a pharmacist's refusal to fill a prescription for an emergency contraception. There are good reasons to promote conscientiousness and to respect such acts of conscience in many, though not all, cases.

Respecting conscientious refusals in health care is an important value, and these refusals should be accommodated unless there are overriding conflicting values. Banning or greatly restricting conscientious refusals in health care could have several negative consequences. It could, according to one analysis, negatively affect the type of people who choose medicine as their vocation and how practicing physicians view and discharge professional responsibilities. It could also foster "callousness" and encourage physicians' "intolerance" of diverse moral beliefs among their patients (and perhaps among their colleagues as well).[42] These possible negative effects are somewhat speculative, but they merit consideration in forming institutional and public policies.

Also meriting consideration is that some conscientious refusals adversely affect patients' and others' legitimate interests in (1) timely access, (2) safe and effective care, (3) respectful care, (4) nondiscriminatory treatment, (5) care that is not unduly burdensome, and (6) privacy and confidentiality. Hence, public policy, professional associations, and health care institutions should seek to recognize and accommodate conscientious refusals as long as they can do so without seriously compromising patients' rights and interests. The metaphor of *balancing* professionals' and patients' rights and interests is commonly used to guide efforts to resolve such conflicts, but it offers only limited guidance and no single model of appropriate response covers all cases.[43]

Institutions such as hospitals and pharmacies can often ensure the timely performance of needed or requested services while allowing conscientious objectors not to perform those services.[44] However, ethical problems arise when, for example, a pharmacist refuses, on grounds of complicity in moral wrongdoing, to transfer a consumer's prescription or to inform the consumer of pharmacies that would fill the prescription. According to one study, only 86% of US physicians surveyed regard themselves as obligated to disclose information about morally controversial medical procedures to patients, and only 71% of US physicians recognize an obligation to refer patients to another physician for such controversial procedures.[45] Consequently, millions of patients in the United States may be under the care of physicians who do not recognize these obligations or are undecided about them.

At a minimum, in our view, health care professionals have an ethical duty to inform prospective employers and prospective patients, clients, and consumers in advance of their personal conscientious objections to performing vital services. Likewise, they have an ethical duty to disclose options for obtaining legal, albeit morally controversial, services; and sometimes they have a duty to provide a referral for those services. They also may have a duty to perform the services in emergency circumstances when the patient is at risk of adverse health effects and a timely referral is not possible.[46]

Determining the appropriate scope of protectable conscientious refusals is a vexing problem, particularly when the refusals involve expansive notions of what counts as assisting or participating in the performance of a personally objectionable action. Such expansive notions sometimes include actions that are only indirectly related to the objectionable procedure. For example, some nurses have claimed conscientious exemption from all forms of participation in the care of patients having an abortion or sterilization, including filling out admission forms or providing post-procedure care. It is often difficult and sometimes impractical for institutions to pursue their mission while exempting objectors to such broadly delineated forms of participation in a procedure.

MORAL IDEALS

We argued in Chapter 1 that norms of obligation in the common morality constitute a moral minimum of requirements that govern everyone. These standards differ from extraordinary moral standards that are not *required* of any person. Moral ideals such as extraordinary generosity are rightly admired and approved by all morally committed persons, and in this respect they are part of the common morality. Extraordinary moral standards come from a morality of aspiration in which individuals, communities, or institutions adopt high ideals not required of others. We can praise and admire those who live up to these ideals, but we cannot blame or criticize persons who do not pursue the ideals.

A straightforward example of a moral ideal in biomedical ethics is found in "expanded access" or "compassionate use" programs that—prior to regulatory approval—authorize access to an investigational drug or device for patients with a serious or immediately life-threatening disease or condition. These patients have exhausted available therapeutic options and are situated so that they cannot participate in a clinical trial of a comparable investigational product. Although it is compassionate and justified to provide some investigational products for therapeutic use, it is generally not obligatory to do so. These programs are compassionate, nonobligatory, and motivated by a goal of providing a good to these patients. The self-imposed moral commitment by the sponsors of the investigational product usually springs from moral ideals of communal service or providing a benefit to individual patients. (See Chapter 6, pp. 224–27, for additional discussion of expanded access programs.)

With the addition of moral ideals, we now have four categories pertaining to moral action: (1) actions that are right and obligatory (e.g., truth-telling); (2) actions that are wrong and prohibited (e.g., murder and rape); (3) actions that are optional and morally neutral, and so neither wrong nor obligatory (e.g., playing chess with a friend); and (4) actions that are optional but morally meritorious and praiseworthy (e.g., sending flowers to a hospitalized friend). We concentrated on the first two in Chapter 1, occasionally mentioning the third. We now focus exclusively on the fourth.

Supererogation and Virtue

Supererogation is a category of moral ideals pertaining principally to ideals of action, but it has important links both to virtues and to Aristotelian ideals of moral excellence.[47] The etymological root of *supererogation* means paying or performing beyond what is owed or, more generally, doing more than is required. This notion has four essential conditions. First, supererogatory acts are optional and neither required nor forbidden by common-morality standards of obligation. Second, supererogatory acts exceed what the common morality of obligation demands, but at least some moral ideals are *endorsed* by all persons committed to the common morality. Third, supererogatory acts are intentionally undertaken to promote the welfare interests of others. Fourth, supererogatory acts are morally good and praiseworthy in themselves and are not merely acts undertaken with good intentions.

Despite the first condition, individuals who act on moral ideals do not always *consider* their actions to be morally optional. Many heroes and saints describe their actions in the language of *ought*, *duty*, and *necessity*: "I had to do it." "I had no choice." "It was my duty." The point of this language is to express a personal sense of obligation, not to state a general obligation. The agent accepts, as a pledge or assignment of personal responsibility, a norm that lays down what ought to be done. At the end of Albert Camus's *The Plague*, Dr. Rieux decides to make a record of those who fought the pestilence. It is to be a record, he says, of "what *had to be done* … despite their personal afflictions, by all who, while unable to be saints but refusing to bow down to pestilences, strive their utmost to be healers."[48] Such healers accept exceptional risks and thereby exceed the obligations of the common morality and of professional associations and traditions.

Many supererogatory acts would be morally obligatory were it not for some abnormal adversity or risk in the face of which the individual elects not to invoke an allowed exemption based on the adversity or risk.[49] If persons have the strength of character that enables them to resist extreme adversity or assume additional risk to fulfill their own conception of their obligations, it makes sense to accept their view that they are under a self-imposed obligation. The hero who says, "I was only doing my duty," is speaking as one who accepts a standard of moral excellence. This hero does not make a mistake in regarding the action as personally required and can view failure as grounds for guilt, although no one else is free to evaluate the act as a moral failure.

Despite the language of "exceptional" and "extreme adversity," not all supererogatory acts are extraordinarily arduous, costly, or risky. Examples of less demanding forms of supererogation include generous gift-giving, volunteering for public service, forgiving another's costly error, and acting from exceptional kindness. Many everyday actions exceed obligation without reaching the highest levels of supererogation. For example, a nurse may put in extra hours of work during the day and return to the hospital at night to visit patients. This nurse's actions are morally excellent, but he or she does not thereby qualify as a saint or hero.

Often we are uncertain whether an action exceeds obligation because the boundaries of obligation and supererogation are ill defined. There may be no clear norm of action, only a virtue of character at work. For example, what is a nurse's role obligation to desperate, terminally ill patients who cling to the nurse for comfort in their few remaining days? If the obligation is that of spending forty hours a week conscientiously fulfilling a job description, the nurse exceeds that obligation by just a few off-duty visits to patients. If the obligation is simply to help patients overcome burdens and meet a series of challenges, a nurse who does so while displaying extraordinary patience, fortitude, and friendliness well exceeds the demands of obligation. Health care professionals sometimes live up to what would ordinarily be a role obligation (such as complying with basic standards of care) while making a sacrifice or taking an additional risk. These cases exceed obligation, but they may not qualify as supererogatory actions.

The Continuum from Obligation to Supererogation

Our analysis may seem to suggest that actions should be classified as either obligatory or beyond the obligatory. The better view, however, is that actions sometimes do not fit neatly into these categories because they fall between the two. Common morality distinctions and ethical theory are not precise enough to determine whether

all actions are morally required or morally elective. This problem is compounded in professional ethics, because professional roles engender obligations that do not bind persons who do not occupy the relevant professional roles. Hence, the two "levels" of the obligatory and the supererogatory lack sharp boundaries both in the common morality and in professional ethics.

Actions may be strictly obligatory, beyond the obligatory, or somewhere between these two classifications. A continuum runs from strict obligation (such as the obligations in the core principles and rules in the common morality) through weaker obligations that are still within the scope of the morally required (such as double-checking one's professional work to be sure that no medical errors have occurred), and on to the domain of the morally nonrequired and the exceptionally virtuous. The nonrequired starts with low-level supererogation, such as walking a visitor lost in a hospital's corridors to a doctor's office. Here an absence of generosity or kindness in helping someone may constitute a small defect in the moral life rather than a failure of obligation. The continuum ends with high-level supererogation, such as heroic acts of self-sacrifice, as in highly risky medical self-experimentation. A continuum exists on each level. The following diagram represents the continuum.

This continuum moves from strict obligation to the most arduous and elective moral ideal. The horizontal line represents a continuum with rough, not sharply defined, breaks. The middle vertical line divides the two general categories but is not meant to indicate a sharp break. Accordingly, the horizontal line expresses a continuum across the four lower categories and expresses the scope of the common morality's reach into the domains of both moral obligations and nonobligatory moral ideals.

Joel Feinberg argues that supererogatory acts are "located on an altogether different scale than obligations."[50] The preceding diagram suggests that this comment is correct in one respect but incorrect in another. The right half of the diagram is not scaled by obligation, whereas the left half is. In this respect, Feinberg's comment is correct. However, the full horizontal line is connected by a single scale of moral value in which the right is continuous with the left. For example, obligatory acts of beneficence and supererogatory acts of beneficence are on the same scale because they are morally of the same kind. The domain of supererogatory ideals is continuous with the domain of norms of obligation by *exceeding* those obligations in accordance with the several defining conditions of supererogation listed previously.

The Place of Ideals in Biomedical Ethics

Many beneficent actions by health care professionals straddle the territory marked in the preceding diagram between *Obligation* and *Beyond Obligation* (in particular, the territory between [2] and [3]). Matters become more complicated when we introduce the distinction discussed in Chapter 1 between professional obligations and obligations incumbent on everyone. Many moral duties established by roles in health care are not moral obligations for persons not in these roles. These duties in medicine and nursing are profession-relative, and some are role obligations even when not formally stated in professional codes. For example, the expectation that physicians and nurses will encourage and cheer despondent patients is a profession-imposed obligation, though not one typically incorporated in a professional code of ethics.

Some customs in the medical community are not well established as obligations, such as the belief that physicians and nurses should efface self-interest and take risks in attending to patients. The nature of "obligations" when caring for patients with SARS (severe acute respiratory syndrome), Ebola, and other diseases with a significant risk of transmission and a significant mortality rate has been controversial, and professional codes and medical association pronouncements have varied.[51] One of the strongest statements of physician duty appeared in the previously mentioned original 1847 Code of Medical Ethics of the American Medical Association (AMA): "when pestilence prevails, it is their [physicians'] duty to face the danger, and to continue their labours for the alleviation of the suffering, even at the jeopardy of their own lives."[52] This statement was retained in subsequent versions of the AMA code until the 1950s, when the statement was eliminated, perhaps in part because of a false sense of the permanent conquest of dangerous contagious diseases.

We usually cannot resolve controversies about duty in face of risk without determining the level of risk—in terms of both the probability and the seriousness of harm—that professionals are expected to assume and setting a threshold beyond which the level of risk is so high that it renders action optional rather than obligatory. The profound difficulty of drawing this line should help us appreciate why some medical associations have urged their members to be courageous and treat patients with potentially lethal infectious diseases, while other associations have advised their members that treatment is optional in many circumstances.[53] Still others have taken the view that both virtue and obligation converge to the conclusion that health care professionals should set aside self-interest, within limits, and that the health care professions should take actions to ensure appropriate care.[54]

Confusion occasionally arises about such matters because of the indeterminate boundaries of what is required in the common morality, what is or should be required in professional communities, and what is a matter of moral character beyond the requirements of moral obligations. In many cases it is doubtful that health care professionals fail to discharge *moral obligations* when they fall short of the highest standards in the profession.

MORAL EXCELLENCE

Aristotelian ethical theory closely connects moral excellence to moral character, moral virtues, and moral ideals. Aristotle succinctly presents this idea: "A truly good and intelligent person … from his resources at any time will do the finest actions he can, just as a good general will make the best use of his forces in war, and a good shoemaker will produce the finest shoe he can from the hides given him, and similarly for all other craftsmen."[55] This passage captures the demanding nature of Aristotle's theory by contrast to ethical theories that focus largely or entirely on the moral minimum of obligations.

The value of this vision of excellence is highlighted by John Rawls, in conjunction with what he calls the "Aristotelian principle":

> The excellences are a condition of human flourishing; they are goods from everyone's point of view. These facts relate them to the conditions of self-respect, and account for their connection with our confidence in our own value. … [T]he virtues are [moral] excellences. … The lack of them will tend to undermine both our self-esteem and the esteem that our associates have for us.[56]

We now draw on this general background in Aristotelian theory and on our prior analysis of moral ideals and supererogation for an account of moral excellence.

The Idea of Moral Excellence

We begin with four considerations that motivate us to examine moral excellence. First, we hope to overcome an undue imbalance in contemporary ethical theory and bioethics that results from focusing narrowly on the moral minimum of obligations while ignoring supererogation and moral ideals.[57] This concentration dilutes the moral life, including our expectations for ourselves, our close associates, and health professionals. If we expect only the moral minimum of obligation, we may lose an ennobling sense of moral excellence. A second and related motivation is our hope to overcome a suppressed skepticism in contemporary ethical theory concerning high ideals in the moral life. Some influential writers note that high moral ideals must compete with other goals and responsibilities in life, and consequently that these ideals can lead persons to neglect other matters worthy of attention, including personal projects, family relationships, friendships, and experiences that broaden outlooks.[58] A third motivation concerns what we call in Chapter 9 the *criterion of comprehensiveness* in an ethical theory. Recognizing the value of moral excellence allows us to incorporate a broad range of moral virtues and forms of supererogation beyond the obligations, rights, and virtues that comprise ordinary morality. Fourth, a model of moral excellence merits pursuit because it indicates what is worthy of aspiration. Morally exemplary lives provide ideals that help guide and inspire us to higher goals and morally better lives.

Aristotelian Ideals of Moral Character

Aristotle maintained that we acquire virtues much as we do skills such as carpentry, playing a musical instrument, and cooking.[59] Both moral and nonmoral skills require training and practice. Obligations play a less central role in his account. Consider, for example, a person who undertakes to expose scientific fraud in an academic institution. It is easy to frame this objective as a matter of obligation, especially if the institution has a policy on fraud. However, suppose this person's correct reports of fraud to superiors are ignored, and eventually her job is in jeopardy and her family receives threats. At some point, she has fulfilled her obligations and is not morally required to pursue the matter further. However, if she does persist, her continued pursuit would be praiseworthy, and her efforts to bring about institutional reform could even reach heroic dimensions. Aristotelian theory could and should frame this situation in terms of the person's level of commitment, the perseverance and endurance shown, the resourcefulness and discernment in marshalling evidence, and the courage as well as the decency and diplomacy displayed in confronting superiors.

An analogy to education illustrates why setting goals beyond the moral minimum is important, especially when discussing moral character. Most of us are trained to aspire to an ideal of education. We are taught to prepare ourselves as best we can. No educational aspirations are too high unless they exceed our abilities and cannot be attained. If we perform at a level below our educational potential, we may consider our achievement a matter of disappointment and regret even if we obtain a university degree. As we fulfill our aspirations, we sometimes expand our goals beyond what we had originally planned. We think of getting another degree, learning another language, or reading widely beyond our specialized training. However, we do not say at this point that we have an *obligation* to achieve at the highest possible level we can achieve.

The Aristotelian model suggests that moral character and moral achievement are functions of self-cultivation and aspiration. Goals of moral excellence can and should enlarge as moral development progresses. Each individual should seek to reach a level as elevated as his or her ability permits, not as a matter of *obligation* but of *aspiration*. Just as persons vary in the quality of their performances in athletics and medical practice, so too in the moral life some persons are more capable than others and deserve more acknowledgment, praise, and admiration. Some persons are sufficiently advanced morally that they exceed what persons less well developed are able to achieve.

Wherever a person is on the continuum of moral development, there will be a goal of excellence that exceeds what he or she has already achieved. This potential to revise our aspirations is centrally important in the moral life. Consider a clinical investigator who uses human subjects in research but who asks only, "What am I obligated to do to protect human subjects?" This investigator's presumption is that once this question has been addressed by reference to a checklist of obligations (for example, government regulations), he or she can ethically proceed with the research. By contrast, in the model we are proposing, this approach is only the starting point. The most important question is, "How could I conduct this research to maximally protect and minimally inconvenience subjects, commensurate with achieving the objectives of the research?" Evading this question indicates that one is morally less committed than one could and probably should be.

The Aristotelian model we have sketched does not expect perfection, only that persons strive toward perfection. This goal might seem impractical, but moral ideals truly can function as practical instruments. As *our* ideals, they motivate us and set out a path that we can climb in stages, with a renewable sense of progress and achievement.

Exceptional Moral Excellence: Saints, Heroes, and Others

Extraordinary persons often function as models of excellence whose examples we aspire to follow. Among the many models, the moral hero and the moral saint are the most celebrated.

The term *saint* has a long history in religious traditions where a person is recognized for exceptional holiness, but, like *hero*, the term *saint* has a secular moral use where a person is recognized for exceptional action or

virtue. Excellence in other-directedness, altruism, and benevolence are prominent features of the moral saint.[60] Saints do their duty and realize moral ideals where most people would fail to do so, and saintliness requires regular fulfillment of duty and realization of ideals over time. It also demands consistency and constancy. We likely cannot make an adequate or final judgment about a person's moral saintliness until the record is complete. By contrast, a person may become a moral hero through a single exceptional action, such as accepting extraordinary risk while discharging duty or realizing ideals. The hero resists fear and the desire for self-preservation in undertaking risky actions that most people would avoid, but the hero also may lack the constancy over a lifetime that distinguishes the saint.

Many who serve as moral models or as persons from whom we draw moral inspiration are not so advanced morally that they qualify as saints or heroes. We learn about good moral character from persons with a limited repertoire of exceptional virtues, such as conscientious health professionals. Consider, for example, John Berger's biography of English physician John Sassall (the pseudonym Berger used for physician John Eskell), who chose to practice medicine in a poverty-ridden, culturally deprived country village in a remote region of northern England. Under the influence of works by Joseph Conrad, Sassall chose this village from an "ideal of service" that reached beyond "the average petty life of self-seeking advancement." Sassall was aware that he would have almost no social life and that the villagers had few resources to pay him, to develop their community, and to attract better medicine, but he focused on their needs rather than his. Progressively, Sassall grew morally as he interacted with members of the community. He developed a deep understanding of, and profound respect for, the villagers. He became a person of exceptional caring, devotion, discernment, conscientiousness, and patience when taking care of the villagers. His moral character deepened year after year. People in the community, in turn, trusted him under adverse and personally difficult circumstances.[61]

From exemplary lives such as that of John Sassall and from our previous analysis, we can extract four criteria of moral excellence.[62] First, Sassall is faithful to a *worthy moral ideal* that he keeps constantly before him in making judgments and performing actions. The ideal is deeply devoted service to a poor and needy community. Second, he has a *motivational structure* that conforms closely to our earlier description of the motivational patterns of virtuous persons who are prepared to forgo certain advantages for themselves in the service of a moral ideal. Third, he has an *exceptional moral character*; that is, he possesses moral virtues that dispose him to perform supererogatory actions of a high order and quality.[63] Fourth, he is a *person of integrity*—both moral integrity and personal integrity—and thus is not overwhelmed by distracting conflicts, self-interest, or personal projects in making judgments and performing actions.

These four conditions are jointly sufficient conditions of *moral excellence*. They are also relevant, but not sufficient, conditions of both moral saintliness and moral heroism. John Sassall does not face extremely difficult tasks, a high level of risk, or deep adversity (although he faces some adversity including his bi-polar condition), and these are typically the sorts of conditions that contribute to making a person a saint or a hero. Exceptional as he is, Sassall is neither a saint nor a hero. To achieve this elevated status, he would have to satisfy additional conditions.

Much admired (though sometimes controversial) examples of moral saints acting from a diverse array of religious commitments are Mahatma Gandhi, Florence Nightingale, Mother Teresa, the 14th Dalai Lama (religious name: Tenzin Gyatso), and Albert Schweitzer. Many examples of moral saints are also found in secular contexts where persons are dedicated to lives of service to the poor and downtrodden. Clear examples are persons motivated to take exceptional risks to rescue strangers.[64] Examples of prominent moral heroes include soldiers, political prisoners, and ambassadors who take substantial risks to save endangered persons by acts such as falling on hand grenades to spare comrades and resisting political tyrants.

Scientists and physicians who experiment on themselves to generate knowledge that may benefit others may be heroes. There are many examples: Daniel Carrion injected blood into his arm from a patient with verruga peruana (an unusual disease marked by many vascular eruptions of the skin and mucous membranes as well as fever and severe rheumatic pains), only to discover that it had given him a fatal disease (Oroya fever). Werner Forssman performed the first heart catheterization on himself, walking to the radiological room with the catheter

sticking into his heart.[65] Daniel Zagury injected himself with an experimental AIDS vaccine, maintaining that his act was "the only ethical line of conduct."[66]

A person can qualify as a moral hero or a moral saint only if he or she meets some combination of the previously listed four conditions of moral excellence. It is too demanding to say that a person must satisfy all four conditions to qualify as a moral hero, but a person must satisfy all four to qualify as a moral saint. This appraisal does not imply that moral saints are more valued or more admirable than moral heroes. We are merely proposing conditions of moral excellence that are more stringent for moral saints than for moral heroes.[67]

To pursue and test this analysis, consider two additional cases.[68] First, reflect on physician David Hilfiker's *Not All of Us Are Saints*, which offers an instructive model of very exceptional but not quite saintly or heroic conduct in his efforts to practice "poverty medicine" in Washington, DC.[69] His decision to leave a rural medical practice in the Midwest to provide medical care to the very poor, including the homeless, reflected both an ambition and a felt obligation. Many health problems he encountered stemmed from an unjust social system, in which his patients had limited access to health care and to other basic social goods that contribute to health. He experienced severe frustration as he encountered major social and institutional barriers to providing poverty medicine, and his patients were often difficult and uncooperative. His frustrations generated stress, depression, and hopelessness, along with vacillating feelings and attitudes including anger, pain, impatience, and guilt. Exhausted by his sense of endless needs and personal limitations, his wellspring of compassion failed to respond one day as he thought it should: "Like those whom on another day I would criticize harshly, I harden myself to the plight of a homeless man and leave him to the inconsistent mercies of the city police and ambulance system. Numbness and cynicism, I suspect, are more often the products of frustrated compassion than of evil intentions."

Hilfiker declared that he is "anything but a saint." He considered the label "saint" to be inappropriate for people, like himself, who have a safety net to protect them. Blaming himself for "selfishness," he redoubled his efforts, but recognized a "gap between who I am and who I would like to be," and he considered that gap "too great to overcome." He abandoned "in frustration the attempt to be Mother Teresa," observing that "there are few Mother Teresas, few Dorothy Days who can give everything to the poor with a radiant joy." Hilfiker did consider many of the people with whom he worked day after day as heroes, in the sense that they "struggle against all odds and survive; people who have been given less than nothing, yet find ways to give."

Second, in *What Really Matters: Living a Moral Life Amidst Uncertainty and Danger*, psychiatrist and anthropologist Arthur Kleinman presents half-a-dozen real-life stories about people who, as the book's subtitle suggests, attempt to live morally in the context of unpredictability and hazard.[70] A story that provided the impetus for his book portrays a woman he names Idi Bosquet-Remarque, a French American who for more than fifteen years was a field representative for several different international aid agencies and foundations, mainly in sub-Saharan Africa. Her humanitarian assistance, carried out almost anonymously, involved working with vulnerable refugees and displaced women and children as well as with the various professionals, public officials, and others who interacted with them. Kleinman presents her as a "moral exemplar," who expressed "our finest impulse to acknowledge the suffering of others and to devote our lives and careers to making a difference (practically and ethically) in their lives, even if that difference must be limited and transient."

At times Bosquet-Remarque was dismayed by various failures, including her own mistakes. She despaired about the value of her work given the overwhelming odds against the people she sought to help, and she recognized some truth in several criticisms of her humanitarian assistance. Faced with daunting obstacles, she persisted because of her deep commitment but eventually experienced physical and emotional burnout, numbness, and demoralization. Nevertheless, she returned to the field because of her deep commitment to her work. Bosquet-Remarque recognized that her motives might be mixed. In addition to her altruism and compassion, she also could have been working out family guilt or seeking to liberate her soul. Despite the ever-present risk of serious injury and even death from violence, she was uncomfortable with the image of the humanitarian worker as "hero."

After Bosquet-Remarque's death in an automobile accident, Kleinman informed her family that he wanted to tell her story. Her mother requested that her daughter not be identified by name: "That way, you will honor what she believed in. Not saints or heroes, but ordinary nameless people doing what they feel they must do, even in extraordinary situations. As a family, we believe in this too."

These observations about ordinary persons who act in extraordinary ways are also relevant to what has been called moral heroism in living organ and tissue donation—a topic to which we now turn.

Living Organ Donation

In light of our moral account thus far, how should we assess a person's offer to donate a kidney to a friend or a stranger?

Health care professionals frequently function as moral gatekeepers to determine who may undertake living donation of organs and tissues for transplantation. Blood donation raises few questions, but in cases of bone marrow donation and the donation of kidneys or portions of livers or lungs, health care professionals must consider whether, when, and from whom to invite, encourage, accept, and effectuate donation. Living organ donation raises challenging ethical issues because the transplant team subjects a healthy person to a variably risky surgical procedure, with no medical benefit to him or her. It is therefore appropriate for transplant teams to probe prospective donors' competence to make such decisions and their understanding, voluntariness, and motives.

Historically, transplant teams were suspicious of living, genetically unrelated donors—particularly of strangers and mere acquaintances but, for a long time, even of emotionally related donors such as spouses and friends. This suspicion had several sources, including concerns about donors' motives and worries about their competence to decide, understanding of the risks, and voluntariness in reaching their decisions. This suspicion increased in cases of nondirected donation, that is, donation not to a particular known individual, but to anyone in need. Such putatively altruistic decisions to donate seemed to require heightened scrutiny. However, in contrast to some professionals' attitudes,[71] a majority of the public in the United States believes that the gift of a kidney to a stranger is reasonable and proper and that, in general, the transplant team should accept it.[72] A key reason is that the offer to donate a kidney whether by a friend, an acquaintance, or a stranger typically does not involve such high risks that serious questions should be triggered about the donor's competence, understanding, voluntariness, or motivation.[73]

Transplant teams can and should decline some heroic offers of organs for moral reasons, even when the donors are competent, their decisions informed and voluntary, and their moral excellence beyond question. For instance, transplant teams have good grounds to decline a mother's offer to donate her heart to save her dying child, because the donation would involve others in directly causing her death. A troublesome case arose when an imprisoned, thirty-eight-year-old father who had already lost one of his kidneys wanted to donate his remaining kidney to his sixteen-year-old daughter whose body had already rejected one kidney transplant.[74] The family insisted that medical professionals and ethics committees had no right to evaluate, let alone reject, the father's act of donation. However, questions arose about the voluntariness of the father's offer (in part because he was in prison), about the risks to him (many patients without kidneys do not thrive on dialysis), about the probable success of the transplant (because of his daughter's problems with her first transplant), and about the costs to the prison system (approximately $40,000 to $50,000 a year for dialysis for the father if he donated the remaining kidney).

We propose that society and health care professionals start with the presumption that living organ donation is praiseworthy but optional. Transplant teams need to subject their criteria for selecting and accepting living donors to public scrutiny to ensure that the teams do not inappropriately use their own values about sacrifice, risk, and the like, as the basis for their judgments.[75] Policies and practices of encouraging prospective living donors are ethically acceptable as long as they do not turn into undue influence or coercion. For instance, it is ethically acceptable to remove financial disincentives for potential donors, such as the costs of post-operative

care, expenses associated with travel and accommodations, and the loss of wages while recovering from donation. It is also ethically acceptable to provide a life insurance policy to reduce risks to the family of the living donor.[76] In the final analysis, live organ donors may not rise to the level of heroes, depending on the risks involved, but many embody a moral excellence that merits society's praise, as well as acceptance by transplant teams in accord with defensible criteria. (In Chapter 9, in each major section, we analyze from several perspectives the case of a father who is reluctant, at least partly because of a lack of courage, to donate a kidney to his dying daughter.)

CONCLUSION

In this chapter we have moved to a moral territory distinct from the principles, rules, obligations, and rights treated in Chapter 1. We have rendered the two domains consistent without assigning priority to one over the other. We have discussed how standards of virtue and character are closely connected to other moral norms, in particular to moral ideals and aspirations of moral excellence that enrich the rights, principles, and rules discussed in Chapter 1. The one domain is not inferior to or derivative from the other, and there is reason to believe that these categories all have a significant place in the common morality.

Still other domains of the moral life of great importance in biomedical ethics remain unaddressed. In Chapter 3 we turn to the chief domain not yet analyzed: moral status.

NOTES

1. 1. For relevant literature on the subjects discussed in Chapter 2 and in the last section of Chapter 9, see Stephen Darwall, ed., *Virtue Ethics* (Oxford: Blackwell, 2003); Roger Crisp and Michael Slote, eds., *Virtue Ethics* (Oxford: Oxford University Press, 1997); Roger Crisp, ed., *How Should One Live? Essays on the Virtues* (Oxford: Oxford University Press, 1996); and Daniel Statman, ed., *Virtue Ethics: A Critical Reader* (Washington, DC: Georgetown University Press, 1997). Many constructive discussions of virtue theory are indebted to Aristotle. For a range of treatments, see Julia Annas, *Intelligent Virtue* (New York: Oxford University Press, 2011) and Annas, "Applying Virtue to Ethics," *Journal of Applied Philosophy* 32 (2015): 1–14; Christine Swanton, *Virtue Ethics: A Pluralistic View* (New York: Oxford University Press, 2003); Nancy Sherman, *The Fabric of Character: Aristotle's Theory of Virtue* (Oxford: Clarendon Press, 1989); Alasdair MacIntyre, *After Virtue: A Study in Moral Theory*, 3rd ed. (Notre Dame, IN: University of Notre Dame Press, 2007) and MacIntyre, *Dependent Rational Animals: Why Human Beings Need the Virtues* (Chicago: Open Court, 1999); Timothy Chappell, ed., *Values and Virtues: Aristotelianism in Contemporary Ethics* (Oxford: Clarendon Press, 2006); and Robert Merrihew Adams, *A Theory of Virtue: Excellence in Being for the Good* (Oxford: Clarendon Press, 2006), and Adams, "A Theory of Virtue: Response to Critics," *Philosophical Studies* 148 (2010): 159–65.
2. 2. Jeremy Bentham, *Deontology or the Science of Morality* (Chestnut Hill, MA: Adamant Media, 2005; reprinted in the Elibron Classics Series of the 1834 edition, originally published in London by Longman et al., 1834), p. 196.
3. 3. This sense of "virtue" is intentionally broad. We do not require, as did Aristotle, that virtue involve habituation rather than a natural character trait. See *Nicomachean Ethics*, trans. Terence Irwin (Indianapolis, IN: Hackett, 1985), 1103^{a}18–19. Nor do we follow St. Thomas Aquinas (relying on a formulation by Peter Lombard), who additionally held that virtue is a good quality of mind by which we live rightly and therefore cannot be put to bad use. See *Treatise on the Virtues* (from *Summa Theologiae*, I–II), Question 55, Arts. 3–4. We treat problems of the definition of "virtue" in more detail in Chapter 9.
4. 4. This definition is the primary use reported in the *Oxford English Dictionary* (*OED*). It is defended philosophically by Alan Gewirth, "Rights and Virtues," *Review of Metaphysics* 38 (1985): 751; and Richard B. Brandt, "The Structure of Virtue," *Midwest Studies in Philosophy* 13 (1988): 76. See also the consequentialist account in Julia Driver, *Uneasy Virtue* (Cambridge: Cambridge University Press, 2001), esp. chap. 4, and Driver, "Response to my Critics," *Utilitas* 16 (2004): 33–41. Edmund Pincoffs presents a definition of virtue in terms of desirable dispositional qualities of persons, in *Quandaries and Virtues:*

Against Reductivism in Ethics (Lawrence: University Press of Kansas, 1986), pp. 9, 73–100. See also MacIntyre, *After Virtue*, chaps. 10–18; and Raanan Gillon, "Ethics Needs Principles," *Journal of Medical Ethics* 29 (2003): 307–12, esp. 309.

5. 5. See the pursuit of this Aristotelian theme in Annas, *Intelligent Virtue*, chap. 5. Elizabeth Anscombe's "Modern Moral Philosophy" (*Philosophy* 33 [1958]: 1–19) is the classic mid-twentieth-century paper on the importance for ethics of categories such as character, virtue, the emotions, and Aristotelian ethics, by contrast to moral theories based on moral law, duty, and principles of obligation.
6. 6. This analysis of practices is influenced by Alasdair MacIntyre, *After Virtue*, esp. chap. 14; and Dorothy Emmet, *Rules, Roles, and Relations* (New York: St. Martin's, 1966). See also Justin Oakley and Dean Cocking, *Virtue Ethics and Professional Roles* (Cambridge: Cambridge University Press, 2001); Oakley, "Virtue Ethics and Bioethics," in *The Cambridge Companion to Virtue Ethics*, ed. Daniel C. Russell (Cambridge: Cambridge University Press, 2013), pp. 197–220; and Tom L. Beauchamp, "Virtue Ethics and Conflict of Interest," in *The Future of Bioethics: International Dialogues*, ed. Akira Akabayashi (Oxford: Oxford University Press, 2014), pp. 688–92.
7. 7. A somewhat similar thesis is defended, in dissimilar ways, in Edmund D. Pellegrino, "Toward a Virtue-Based Normative Ethics for the Health Professions," *Kennedy Institute Ethics Journal* 5 (1995): 253–77. See also John Cottingham, "Medicine, Virtues and Consequences," in *Human Lives: Critical Essays on Consequentialist Bioethics*, ed. David S. Oderberg (New York: Macmillan, 1997); Alan E. Armstrong, *Nursing Ethics: A Virtue-Based Approach* (New York: Palgrave Macmillan, 2007); and Jennifer Radden and John Z. Sadler, *The Virtuous Psychiatrist: Character Ethics in Psychiatric Practice* (New York: Oxford University Press, 2010).
8. 8. Charles L. Bosk, *Forgive and Remember: Managing Medical Failure*, 2nd ed. (Chicago: University of Chicago Press, 2003). In addition to the three types of error we mention, Bosk recognizes a fourth type: "quasi-normative errors," based on the attending's special protocols. In the Preface to the second edition, he notes that his original book did not stress as much as it should have the problems that were created when normative and quasi-normative breaches were treated in a unitary fashion (p. xxi).
9. 9. Thomas Percival, *Medical Ethics; or a Code of Institutes and Precepts, Adapted to the Professional Conduct of Physicians and Surgeons* (Manchester, UK: S. Russell, 1803), pp. 165–66. This book formed the substantive basis of the first American Medical Association code in 1847.
10. 10. For this shift, see Gerald R. Winslow, "From Loyalty to Advocacy: A New Metaphor for Nursing," *Hastings Center Rep*ort 14 (June 1984): 32–40; and Helga Kuhse, *Caring: Nurses, Women and Ethics* (Oxford, UK: Blackwell, 1997), esp. chaps. 1, 2, and 9.
11. 11. See the virtue-based approach to nursing ethics in Armstrong, *Nursing Ethics: A Virtue-Based Approach.*
12. 12. Contrast Virginia Held's argument for a sharp distinction between the ethics of care and virtue ethics on the grounds that the former focuses on relationships and the latter on individuals' dispositions: *The Ethics of Care: Personal, Political, and Global* (New York: Oxford University Press, 2006). We are skeptical of her argument, and of the similar view developed by Nel Noddings in "Care Ethics and Virtue Ethics," in *The Routledge Companion to Virtue Ethics*, ed., Lorraine Besser-Jones and Michael Slote (London: Routledge, 2015), pp. 401–14. Drawing on related themes, Ruth Groenhout challenges the standard taxonomies that lump a feminist ethic of care together with virtue ethics (developed from a nonfeminist history); see her "Virtue and a Feminist Ethic of Care," in *Virtues and Their Vices*, ed. Kevin Timpe and Craig A. Boyd (Oxford: Oxford University Press, 2014), pp. 481–501. For an argument closer to ours, see Raja Halwani, "Care Ethics and Virtue Ethics," *Hypatia* 18 (2003): 161–92.
13. 13. Carol Gilligan, *In a Different Voice* (Cambridge, MA: Harvard University Press, 1982), esp. p. 21. See also her "Mapping the Moral Domain: New Images of Self in Relationship," *Cross Currents* 39 (Spring 1989): 50–63.
14. 14. Gilligan and others deny that the two distinct voices correlate strictly with gender. See Gilligan and Susan Pollak, "The Vulnerable and Invulnerable Physician," in *Mapping the Moral Domain*, ed. C. Gilligan, J. Ward, and J. Taylor (Cambridge, MA: Harvard University Press, 1988), pp. 245–62.
15. 15. See Gilligan and G. Wiggins, "The Origins of Morality in Early Childhood Relationships," in *The Emergence of Morality in Young Children*, ed. J. Kagan and S. Lamm (Chicago: University of Chicago Press, 1988). See also Margaret Olivia Little, "Care: From Theory to Orientation and Back," *Journal of Medicine and Philosophy* 23 (1998): 190–209.

16. 16. Our formulation of these criticisms is influenced by Alisa L. Carse, "The 'Voice of Care': Implications for Bioethical Education," *Journal of Medicine and Philosophy* 16 (1991): 5–28, esp. 8–17. For assessment of such criticisms, see Abraham Rudnick, "A Meta-Ethical Critique of Care Ethics," *Theoretical Medicine* 22 (2001): 505–17.
17. 17. Alisa L. Carse, "Impartial Principle and Moral Context: Securing a Place for the Particular in Ethical Theory," *Journal of Medicine and Philosophy* 23 (1998): 153–69.
18. 18. See Christine Grady and Anthony S. Fauci, "The Role of the Virtuous Investigator in Protecting Human Research Subjects," *Perspectives in Biology and Medicin*e 59 (2016): 122–31; Nel Noddings, *Caring: A Feminine Approach to Ethics and Moral Education*, 2nd ed. (Berkeley: University of California Press, 2003), and the evaluation of Noddings's work in Halwani, "Care Ethics and Virtue Ethics," esp. pp. 162ff.
19. 19. See Nancy Sherman, *The Fabric of Character*, pp. 13–55; and Martha Nussbaum, *Love's Knowledge* (Oxford: Oxford University Press, 1990). On "attention" in medical care, see Margaret E. Mohrmann, *Attending Children: A Doctor's Education* (Washington, DC: Georgetown University Press, 2005).
20. 20. Carse, "The 'Voice of Care,'" p. 17.
21. 21. Other virtues are similarly important. We treat several later in this chapter and in Chapter 9. On the historical role of a somewhat different collection of central virtues in medical ethics and their connection to vices, especially since the eighteenth century, see Frank A. Chervenak and Laurence B. McCullough, "The Moral Foundation of Medical Leadership: The Professional Virtues of the Physician as Fiduciary of the Patient," *American Journal of Obstetrics and Gynecology* 184 (2001): 875–80.
22. 22. Edmund D. Pellegrino, "Toward a Virtue-Based Normative Ethics," p. 269. Compassion is often regarded as one of the major marks of an exemplary health care professional. See Helen Meldrum, *Characteristics of Compassion: Portraits of Exemplary Physicians* (Sudbury, MA; Jones and Bartlett, 2010).
23. 23. See Lawrence Blum, "Compassion," in *Explaining Emotions*, ed. Amélie Oksenberg Rorty (Berkeley: University of California Press, 1980); and David Hume, *A Dissertation on the Passions*, ed. Tom L. Beauchamp (Oxford: Clarendon Press, 2007), Sect. 3, §§ 4–5.
24. 24. Martha Nussbaum, *Upheavals of Thought: The Intelligence of Emotions* (Cambridge: Cambridge University Press, 2001), p. 302. Part II of this book is devoted to compassion.
25. 25. See Jodi Halpern, *From Detached Concern to Empathy: Humanizing Medical Practice* (New York: Oxford University Press, 2001). For a variety of largely positive essays on empathy, see Howard Spiro et al., eds., *Empathy and the Practice of Medicine* (New Haven, CT: Yale University Press, 1993); and Ellen Singer More and Maureen A. Milligan, eds., *The Empathic Practitioner: Empathy, Gender, and Medicine* (New Brunswick, NJ: Rutgers University Press, 1994). A valuable set of philosophical and psychological perspectives on empathy appears in Amy Coplan and Peter Goldie, eds., *Empathy: Philosophical and Psychological Perspectives* (Oxford: Oxford University Press, 2011). Jean Decety, ed., *Empathy: From Bench to Bedside* (Cambridge, MA: MIT Press, 2012) includes several essays in Part VI on "Empathy in Clinical Practice." For dangers of an overemphasis on empathy in medicine, see Jane Mcnaughton, "The Art of Medicine: The Dangerous Practice of Empathy," *Lancet* 373 (2009): 1940–1941. Paul Bloom offers a sustained psychological argument against empathy in favor of "rational compassion" in health care, and many other areas, in his *Against Empathy: The Case for Rational Compassion* (New York: Ecco Press of HarperCollins, 2016). Some commentators on his thesis recognize the legitimacy of his concerns, for instance, about empathy in health care, but call for a more nuanced perspective and greater appreciation of the value of empathy. See the discussion in response to his essay entitled "Against Empathy" in a Forum in the *Boston Review*, September 10, 2014, available at http://bostonreview.net/forum/paul-bloom-against-empathy (accessed July 22, 2018). Much in this debate hinges on different interpretations of the concept, criteria, and descriptions of empathy.
26. 26. David Hume, *A Treatise of Human Nature*, ed. David Fate Norton and Mary Norton (Oxford: Clarendon Press, 2007), 3.3.1.7.
27. 27. Baruch Brody, "Case No. 25. 'Who Is the Patient, Anyway': The Difficulties of Compassion," in *Life and Death Decision Making* (New York: Oxford University Press, 1988), pp. 185–88.
28. 28. Aristotle, *Nicomachean Ethics*, trans. Terence Irwin, 2nd ed. (Indianapolis: Hackett, 2000), $1106^{b}15$–29, $1141^{a}15$–$1144^{b}17$.

29. 29. Annette Baier, "Trust, Suffering, and the Aesculapian Virtues," in *Working Virtue: Virtue Ethics and Contemporary Moral Problems*, ed. Rebecca L. Walker and Philip J. Ivanhoe (Oxford: Clarendon Press, 2007), p. 137.
30. 30. See Annette Baier's "Trust and Antitrust" and two later essays on trust in her *Moral Prejudices* (Cambridge, MA: Harvard University Press, 1994); Nancy N. Potter, *How Can I Be Trusted: A Virtue Theory of Trustworthiness* (Lanham, MD: Rowman & Littlefield, 2002); Philip Pettit, "The Cunning of Trust," *Philosophy & Public Affairs* 24 (1995): 202–25; and Pellegrino and Thomasma, *The Virtues in Medical Practice*, chap. 5.
31. 31. Aristotle, *Eudemian Ethics*, $1242^{b}23–1243^{a}13$, in *The Complete Works of Aristotle*, ed. Jonathan Barnes (Princeton, NJ: Princeton University Press, 1984).
32. 32. For discussions of the erosion of trust in medicine, see Robert J. Blendon, John M. Benson, and Joachim O. Hero, "Public Trust in Physicians—U.S. Medicine in International Perspective" (a project studying 29 industrialized countries sponsored by the Robert Wood Johnson Foundation), *New England Journal of Medicine* 371 (2014): 1570–72; David A. Axelrod and Susan Dorr Goold, "Maintaining Trust in the Surgeon-Patient Relationship: Challenges for the New Millennium," *Archives of Surgery* 135 (January 2000), available at https://jamanetwork.com/journals/jamasurgery/fullarticle/390488 (accessed March 17, 2018); David Mechanic, "Public Trust and Initiatives for New Health Care Partnerships," *Milbank Quarterly* 76 (1998): 281–302; Pellegrino and Thomasma in *The Virtues in Medical Practice*, pp. 71–77; and Mark A. Hall, "The Ethics and Empirics of Trust," in *The Ethics of Managed Care: Professional Integrity and Patient Rights*, ed. W. B. Bondeson and J. W. Jones (Dordrecht, Netherlands: Kluwer, 2002), pp. 109–26. Broader explorations of trustworthiness, trust, and distrust appear in Russell Hardin's *Trust and Trustworthiness,* Russell Sage Foundation Series on Trust, vol. 4 (New York: Russell Sage Foundation Publications, 2004). See further Onora O'Neill's proposals to restore trust in medical and other contexts where mistrust results from factors such as bureaucratic structures of accountability, excessive transparency, and public culture: *A Question of Trust* (Cambridge: Cambridge University Press, 2002) and *Autonomy and Trust in Bioethics* (Cambridge: Cambridge University Press, 2003).
33. 33. Brody, *Life and Death Decision Making*, p. 35. On the interpretation of integrity as a virtue, see Damian Cox, Marguerite La Caze, and Michael Levine, "Integrity," *The Stanford Encyclopedia of Philosophy* (Spring 2017 Edition), ed. Edward N. Zalta, available at https://plato.stanford.edu/archives/spr2017/entries/integrity/ (accessed March 27, 2018).
34. 34. On the connection of, and the distinction between, autonomy and integrity, see Carolyn McLeod, "How to Distinguish Autonomy from Integrity," *Canadian Journal of Philosophy* 35 (2005): 107–33.
35. 35. On integrity as a virtue in the medical professions, see Edmund D. Pellegrino, "Codes, Virtue, and Professionalism," in *Methods of Medical Ethics*, ed. Jeremy Sugarman and Daniel P. Sulmasy, revised 2nd ed. (Washington, DC: Georgetown University Press, 2010), pp. 91–107, esp. 94; and Michael Wreen, "Medical Futility and Physician Discretion," *Journal of Medical Ethics* 30 (2004): 275–78.
36. 36. For useful discussions of this question in nursing, see Martin Benjamin and Joy Curtis, *Ethics in Nursing: Cases, Principles, and Reasoning*, 4th ed. (New York: Oxford University Press, 2010), pp. 122–26; and Betty J. Winslow and Gerald Winslow, "Integrity and Compromise in Nursing Ethics," *Journal of Medicine and Philosophy* 16 (1991): 307–23. A wide-ranging discussion is found in Martin Benjamin, *Splitting the Difference: Compromise and Integrity in Ethics and Politics* (Lawrence: University Press of Kansas, 1990).
37. 37. For a historically grounded critique of such conceptions and a defense of conscience as a virtue, see Douglas C. Langston, *Conscience and Other Virtues: From Bonaventure to MacIntyre* (University Park: Pennsylvania State University Press, 2001). For another historical perspective, see Richard Sorabji, *Moral Conscience Through the Ages: Fifth Century BCE to the Present* (Chicago: University of Chicago Press, 2014).
38. 38. Bernard Williams, "A Critique of Utilitarianism," in J. J. C. Smart and Williams, *Utilitarianism: For and Against* (Cambridge: Cambridge University Press, 1973), pp. 97–98.
39. 39. We here draw from two sources: Hannah Arendt, *Crises of the Republic* (New York: Harcourt, Brace, Jovanovich, 1972), p. 62; and John Stuart Mill, *Utilitarianism*, chap. 3, pp. 228–29, and *On Liberty*, chap. 3, p. 263, in *Collected Works of John Stuart Mill*, vols. 10, 18 (Toronto, Canada: University of Toronto Press, 1969, 1977).

40. 40. Carl H. Fellner, "Organ Donation: For Whose Sake?" *Annals of Internal Medicine* 79 (October 1973): 591.
41. 41. See James F. Childress, "Appeals to Conscience," *Ethics* 89 (1979): 315–35; Larry May, "On Conscience," *American Philosophical Quarterly* 20 (1983): 57–67; and C. D. Broad, "Conscience and Conscientious Action," in *Moral Concepts*, ed. Joel Feinberg (Oxford: Oxford University Press, 1970), pp. 74–79. See also Daniel P. Sulmasy, "What Is Conscience and Why Is Respect for It So Important?" *Theoretical Medicine and Bioethics* 29 (2008): 135–49; and Damian Cox, Marguerite La Caze, and Michael Levine, "Integrity," *The Stanford Encyclopedia of Philosophy* (Spring 2017 Edition), ed. Edward N. Zalta, available at https://plato.stanford.edu/archives/spr2017/entries/integrity/ (accessed February 25, 2018).
42. 42. Douglas B. White and Baruch Brody, "Would Accommodating Some Conscientious Objections by Physicians Promote Quality in Medical Care?" *JAMA* 305 (May 4, 2011): 1804–5.
43. 43. For several models, see Rebecca Dresser, "Professionals, Conformity, and Conscience," *Hastings Center Report* 35 (November–December 2005): 9–10; Mark R. Wicclair, *Conscientious Objection in Health Care: An Ethical Analysis* (Cambridge: Cambridge University Press, 2011); Alta R. Charo, "The Celestial Fire of Conscience—Refusing to Deliver Medical Care," *New England Journal of Medicine* 352 (2005): 2471–73; and Elizabeth Fenton and Loren Lomasky, "Dispensing with Liberty: Conscientious Refusal and the 'Morning-After Pill,'" *Journal of Medicine and Philosophy* 30 (2005): 579–92.
44. 44. See Holly Fernandez Lynch, *Conflicts of Conscience: An Institutional Compromise* (Cambridge, MA: MIT Press, 2008).
45. 45. The rest of the physicians are opposed or undecided. Farr A. Curlin et al., "Religion, Conscience, and Controversial Clinical Practices," *New England Journal of Medicine* 356 (February 8, 2007): 593–600.
46. 46. Dan W. Brock offers a similar framework for ethical analysis in what he calls the "conventional compromise" in "Conscientious Refusal by Physicians and Pharmacists: Who Is Obligated to Do What, and Why?" *Theoretical Medicine and Bioethics* 29 (2008): 187–200. For the legal framework in the United States, see Elizabeth Sepper, "Conscientious Refusals of Care," in *The Oxford Handbook of U.S. Health Law*, ed. I. Glenn Cohen, Allison Hoffman, and William M. Sage (New York: Oxford University Press, 2017), chap. 16.
47. 47. Our analysis is indebted to David Heyd, *Supererogation: Its Status in Ethical Theory* (Cambridge: Cambridge University Press, 1982); Heyd, "Tact: Sense, Sensitivity, and Virtue," *Inquiry* 38 (1995): 217–31; Heyd, "Obligation and Supererogation," *Encyclopedia of Bioethics*, 3rd ed. (New York: Thomson Gale, 2004), vol. 4, pp. 1915–20; and Heyd, "Supererogation," *The Stanford Encyclopedia of Philosophy* (Spring 2016 Edition), ed. Edward N. Zalta, available at https://plato.stanford.edu/archives/spr2016/entries/supererogation (accessed March 27, 2018). We are also indebted to J. O. Urmson, "Saints and Heroes," *Essays in Moral Philosophy*, ed. A. I. Melden (Seattle: University of Washington Press, 1958), pp. 198–216; John Rawls, *A Theory of Justice* (Cambridge, MA: Harvard University Press, 1971; rev. ed. 1999), pp. 116–17, 438–39, 479–85 (1999: 100–101, 385–86, 420–25); Joel Feinberg, "Supererogation and Rules," *Ethics* 71 (1961); and Gregory Mellema, *Beyond the Call of Duty: Supererogation, Obligation, and Offence* (Albany: State University of New York Press, 1991). For central connections between virtue and supererogation, see Roger Crisp, "Supererogation and Virtue," in *Oxford Studies in Normative Ethics* (vol. 3), ed. Mark Timmons (Oxford: Oxford University Press, 2013), article 1.
48. 48. Albert Camus, *The Plague*, trans. Stuart Gilbert (New York: Knopf, 1988), p. 278. Italics added.
49. 49. The formulation in this sentence relies in part on Rawls, *A Theory of Justice*, p. 117 (1999 edition, p. 100).
50. 50. Feinberg, "Supererogation and Rules," 397.
51. 51. See Dena Hsin-Chen and Darryl Macer, "Heroes of SARS: Professional Roles and Ethics of Health Care Workers," *Journal of Infection* 49 (2004): 210–15; Joseph J. Fins, "Distinguishing Professionalism and Heroism When Disaster Strikes: Reflections on 9/11, Ebola, and Other Emergencies," *Cambridge Quarterly of Healthcare Ethics* 24 (October 2015): 373–84; Angus Dawson, "Professional, Civic, and Personal Obligations in Public Health Emergency Planning and Response," in *Emergency Ethics: Public Health Preparedness and Response*, ed. Bruce Jennings, John D. Arras, Drue H. Barrett, and Barbara A. Ellis (New York: Oxford University Press, 2016), pp. 186–219. Early discussions of HIV/AIDS, when there were major concerns about transmission in the clinical setting, frequently addressed the clinician's

responsibility to treat. Examples include Bernard Lo, "Obligations to Care for Persons with Human Immunodeficiency Virus," *Issues in Law & Medicine* 4 (1988): 367–81; Doran Smolkin, "HIV Infection, Risk Taking, and the Duty to Treat," *Journal of Medicine and Philosophy* 22 (1997): 55–74; and John Arras, "The Fragile Web of Responsibility: AIDS and the Duty to Treat," *Hastings Center Report* 18 (April–May 1988): S10–20.

52. 52. American Medical Association (AMA), *Code of Medical Ethics of the American Medical Association*, adopted May 1847 (Philadelphia: T.K. and P.G. Collins, 1848), available at http://ethics.iit.edu/ecodes/sites/default/files/Americaan%20Medical%20Association%20Code%20of%20Medical%20Ethics%20%281847%29.pdf (accessed March 17, 2018).
53. 53. See American Medical Association, Council on Ethical and Judicial Affairs, "Ethical Issues Involved in the Growing AIDS Crisis," *Journal of the American Medical Association* 259 (March 4, 1988): 1360–61.
54. 54. Health and Public Policy Committee, American College of Physicians and Infectious Diseases Society of America, "The Acquired Immunodeficiency Syndrome (AIDS) and Infection with the Human Immunodeficiency Virus (HIV)," *Annals of Internal Medicine* 108 (1988): 460–61. See further Edmund D. Pellegrino, "Character, Virtue, and Self-Interest in the Ethics of the Professions," *Journal of Contemporary Health Law and Policy* 5 (1989): 53–73, esp. 70–71.
55. 55. Aristotle, *Nicomachean Ethics*, trans. Irwin, $1101^{a}1$–7.
56. 56. Rawls, *A Theory of Justice*, pp. 443–45 (1999 edition: 389–91). On the Aristotelian principle, see pp. 424–33 (1999 edition: 372–80).
57. 57. Urmson recognized this problem in "Saints and Heroes," pp. 206, 214. Imbalance is found in forms of utilitarianism that make strong demands of obligation. However, see the attempt to revise consequentialism to bring it in line with common moral intuitions in Douglas W. Portman, "Position-Relative Consequentialism, Agent-Centered Options, and Supererogation," *Ethics* 113 (2003): 303–32.
58. 58. A reasonable skepticism is evident in some influential philosophical works such as those of Susan Wolf (in the article cited below), Philippa Foot, Bernard Williams, and Thomas Nagel.
59. 59. Aristotle, *Nicomachean Ethics*, trans. Irwin, $1103^{a}32$–$1103^{b}1$.
60. 60. Edith Wyschogrod offers a definition of a "saintly life" as "one in which compassion for the other, irrespective of cost to the saint, is the primary trait." Wyschogrod, *Saints and Postmodernism: Revisioning Moral Philosophy* (Chicago: University of Chicago Press, 1990), pp. xiii, xxii, et passim.
61. 61. John Berger (and Jean Mohr, photographer), *A Fortunate Man: The Story of a Country Doctor* (London: Allen Lane, the Penguin Press, 1967), esp. pp. 48, 74, 82ff, 93ff, 123–25, 135. Lawrence Blum pointed us to this book and influenced our perspective on it. Sassall's wife played a critical role in running his medical practice and helping him deal with his manic-depressive illness; she receives little attention in the book, which is, however, dedicated to her. She died in 1981, and he committed suicide the next year. See Roger Jones, "Review: *A Fortunate Man*," *British Journal of General Practice*, February 9, 2015, available at http://bjgplife.com/2015/02/09/review-a-fortunate-man/ (accessed July 20, 2018). See also Gavin Francis, "John Berger's *A Fortunate Man*: A Masterpiece of Witness," *Guardian*, February 7, 2015, available at https://www.theguardian.com/books/2015/feb/07/john-sassall-country-doctor-a-fortunate-man-john-berger-jean-mohr (accessed, July 20, 2018).
62. 62. Our conditions of moral excellence are indebted to Lawrence Blum, "Moral Exemplars," *Midwest Studies in Philosophy* 13 (1988): 204. See also Blum's "Community and Virtue," in *How Should One Live?: Essays on the Virtues*, ed. Crisp.
63. 63. Our second and third conditions are influenced by the characterization of a saint in Susan Wolf's "Moral Saints," *Journal of Philosophy* 79 (1982): 419–39. For a pertinent critique of Wolf's interpretation, see Robert Merrihew Adams, "Saints," *Journal of Philosophy* 81 (1984), reprinted in Adams, *The Virtue of Faith and Other Essays in Philosophical Theology* (New York: Oxford University Press, 1987), pp. 164–73.
64. 64. For an examination of some twenty-first-century figures who lived under extreme conditions with exceptional moral commitment, see Larissa MacFarquhar, *Strangers Drowning: Impossible Idealism, Drastic Choices, and the Urge to Help* (New York: Penguin Books, 2016).
65. 65. Jay Katz, ed., *Experimentation with Human Beings* (New York: Russell Sage Foundation, 1972), pp. 136–40; Lawrence K. Altman, *Who Goes First? The Story of Self-Experimentation in Medicine*, 2nd ed., with a new preface (Berkeley: University of California Press, 1998), pp. 1–5, 39–50, et passim.

66. 66. Philip J. Hilts, "French Doctor Testing AIDS Vaccine on Self," *Washington Post*, March 10, 1987, p. A7; Altman, *Who Goes First?*, pp. 26–28.
67. 67. We will not consider whether these conditions point to a still higher form of moral excellence: the combination of saint and hero in one person. There have been such extraordinary persons, and we could make a case that some of these extraordinary figures are more excellent than others. But at this level of moral exemplariness, such fine distinctions serve no purpose.
68. 68. These cases can be read as suggesting that many people who are commonly called heroes or saints are not very different from good and decent but morally ordinary people. This theory is not explored here (except implicitly in our account of the continuum from ordinary morality to supererogation), but it is examined in Andrew Michael Flescher, *Heroes, Saints, and Ordinary Morality* (Washington: Georgetown University Press, 2003). Flescher provides historical examples of people commonly regarded as saints or heroes.
69. 69. David Hilfiker, *Not All of Us Are Saints: A Doctor's Journey with the Poor* (New York: Hill & Wang, 1994). The summaries and quotations that follow come from this book. His earlier book, *Healing the Wounds: A Physician Looks at His Work* (New York: Pantheon, 1985) focuses on his previous experiences as a family physician in rural Minnesota. The personal problems he (and some others we discuss) faced underline a critical point in this chapter: difficulties that can arise in balancing a commitment to a moral ideal or moral excellence with personal needs.
70. 70. Arthur Kleinman, *What Really Matters: Living a Moral Life Amidst Uncertainty and Danger* (New York: Oxford University Press, 2006), chap. 3. The quotations are from this work.
71. 71. For the attitudes of nephrologists, transplant nephrologists, transplant surgeons, and the like, see Carol L. Beasley, Alan R. Hull, and J. Thomas Rosenthal, "Living Kidney Donation: A Survey of Professional Attitudes and Practices," *American Journal of Kidney Diseases* 30 (October 1997): 549–57; and Reginald Y. Gohh, Paul E. Morrissey, Peter N. Madras, et al., "Controversies in Organ Donation: The Altruistic Living Donor," *Nephrology Dialysis Transplantation* 16 (2001): 619–21, available at https://academic.oup.com/ndt/article/16/3/619/1823109 (accessed February 26, 2018). Even though strong support now exists for living kidney donation, actual medical practice is not uniformly in agreement.
72. 72. See Aaron Spital and Max Spital, "Living Kidney Donation: Attitudes Outside the Transplant Center," *Archives of Internal Medicine* 148 (May 1988): 1077–80; Aaron Spital, "Public Attitudes toward Kidney Donation by Friends and Altruistic Strangers in the United States," *Transplantation* 71 (2001): 1061–64.
73. 73. From 1996 to 2005, as living kidney donation overall doubled in the United States, the annual percentage of genetically unrelated kidney donors (excluding spouses) rose from 5.9% to 22%. *2006 Annual Report of the U.S. Organ Procurement and Transplantation Network and the Scientific Registry of Transplant Recipients: Transplant Data 1996–2005* (Rockville, MD: Health Resources and Services Administration, Healthcare Systems Bureau, Division of Transplantation, 2006). During the years 2001–3, acts of living organ donation outnumbered acts of deceased organ donation, but living organ donation, which had increased for the preceding five years, declined steadily after 2004 for both kidneys and livers. See A. S. Klein, E. E. Messersmith, L. E. Ratner, et al., "Organ Donation and Utilization in the United States, 1999–2008," *American Journal of Transplantation* 10 (Part 2) (2010): 973–86. This slide has continued. See James R. Rodrigue, Jesse D. Schold, and Didier A. Mandelbrot, "The Decline in Living Kidney Donation in the United States: Random Variation or Cause for Concern?" *Transplantation* 96 (2013): 767–73.
74. 74. Evelyn Nieves, "Girl Awaits Father's 2nd Kidney, and Decision by Medical Ethicists," *New York Times*, December 5, 1999, pp. A1, A11.
75. 75. See Linda Wright, Karen Faith, Robert Richardson, and David Grant, "Ethical Guidelines for the Evaluation of Living Organ Donors," *Canadian Journal of Surgery* 47 (December 2004): 408–12. See also A. Tong, J. R. Chapman, G. Wong, et al., "Living Kidney Donor Assessment: Challenges, Uncertainties and Controversies among Transplant Nephrologists and Surgeons," *American Journal of Transplantation* 13 (2013): 2912–23. For further examination of ethical issues in living organ donation, see James F. Childress and Cathryn T. Liverman, eds., *Organ Donation: Opportunities for Action* (Washington, DC: National Academies Press, 2006), chap. 9.
76. 76. A vigorous debate continues about whether it would be ethically acceptable to add financial incentives for living organ donation, beyond removing financial disincentives. Such incentives would change some donors' motivations for donation, which already may include factors in addition to their altruism.

3

Moral Status

The previous two chapters concentrated on moral agents and their obligations, rights, and virtues. Little consideration has been given to whom the obligations are owed, why we have obligations to some beings and not others, and which beings have rights and which do not. This chapter is devoted to these questions of moral status, also referred to as moral standing and moral considerability.[1]

The terms *status* and *standing* have been transported to ethics from the notion of legal standing. In a weak sense, "moral status" refers to a position, grade, or rank of moral importance. In a strong sense, "moral status" means to have rights or the functional equivalent of rights. Any being has moral status if moral agents have moral obligations to it, the being has welfare interests, and the moral obligations owed to it are based on its interests.[2]

The Problem of Moral Status

The problem of moral status begins with questions about which entities, individuals, and groups are protected by moral norms. For example, what should we say about human embryonic stem cells? Human eggs? Embryos? Fetuses? Newborn infants? Anencephalic babies? The mentally disabled? Persons who are unable to distinguish right from wrong? The seriously demented? Those incurring a permanent loss of consciousness? The brain-dead? Cadavers? Nonhuman animals used in medical research? A biologically modified animal designed to carry a human fetus to term? Chimeric animals, transgenic animals, and other new life forms created in research? Do the members of each of these groups deserve moral protections or have moral rights? If so, do they deserve the same complement of protections and rights afforded to competent adult humans?[3]

Throughout much of human history, collections of human beings such as racial groupings, tribes, enemies in war, and effectively all nonhuman animals have been treated as less than persons. Accordingly, they were assigned either no moral status or a low-level of moral status and were accorded no moral rights (historically, slaves in many societies) or fewer or weaker rights (historically, women in many societies).[4] Still common, though controversial, presumptions in medicine and biomedical ethics indicate that some groups have no moral rights (e.g., animals used in biomedical research) and that some groups have fewer or weaker rights (e.g., human embryos used in research).

Surrogate decision making also raises questions about moral status. When a once competent person is deemed incompetent and needs a surrogate decision maker, the person does not lose all moral protections and forms of moral respect. Many obligations to these individuals continue, and some new obligations may arise. Nonetheless, the recognition of a surrogate as the rightful decision maker entails that the incompetent individual has lost some rights of decision making, and in this respect the individual's moral status is lower than it previously was. Any "decision" that such an individual might make (e.g., to leave a nursing home) does not have the same moral authority it had prior to the determination of incompetency. At least some of our obligations to the person have shifted and some have ceased. For example, we may no longer be obligated to obtain first-party informed consent from this individual, in which case consent must be obtained from a surrogate decision maker. The criterion of mental incompetence is one among many commonly employed in assessing moral status and in determining rights and obligations.

Similar questions arise about what we owe to small children when we involve them in pediatric research that holds out no promise of direct benefit for child subjects because the goal of the research is to develop new treatments for children in the future. We often assert that we owe vulnerable parties more, not fewer, protections. Yet children involved in research that is not intended to benefit them have sometimes been treated as if they have a diminished moral status.

Another example of problems of moral status comes from cases of pregnant women who are brain-dead but whose biological capacities are artificially maintained for several weeks to enable the fetus they are carrying to be born.[5] Ordinarily, we do not think of dead people as having a moral status that affords them a right to be kept biologically functioning. Moreover, maintaining a brain-dead pregnant woman's body against her formerly stated wishes implies that she has been categorized as having a *lower* moral status than other corpses because her body is subjected to extreme measures—sometimes for months—to benefit the fetus, the woman's partner, or the next of kin in the family.[6]

The central ethical question is whether a fetus has rights stronger than those of a brain-dead pregnant woman whose advance directive expresses her wish to stop all technology at the point of brain death. Beliefs about the moral status of the fetus are powerful motivating considerations in some cases, but the fetus is not the only individual with moral status and rights at the point of the pregnant woman's brain death. Discussion continues about whether a brain-dead woman in this situation has rights that can legitimately be asserted in her advance directive and whether maintaining her body to sustain the pregnancy violates those rights.[7]

Finally, views of and practices toward the many nonhuman animals that we use in biomedical research raise moral status questions. At times we appear to treat them primarily as utilitarian means to the ends of science, facilitated by the decisions of some person or group considered to be their stewards. The implication is that laboratory animals are not morally protected against invasive, painful, and harmful forms of experimentation, and perhaps that they lack moral status altogether. An outright denial of moral status is implausible in light of the fact that virtually every nation and major scientific association has guidelines to alleviate, diminish, or otherwise limit what can be done to animals in biomedical research. It is today generally accepted that animals used in research have some level of moral status, though it often remains unclear which moral considerations warrant this judgment.

At the root of these questions is a rich body of theoretical issues and practical problems about moral status.

THEORIES OF MORAL STATUS

To have moral status is to deserve at least some of the protections afforded by moral norms, including the principles, rules, obligations, and rights discussed in Chapter 1. These protections are afforded only to entities that can be morally wronged by actions. Here is a simple example: We wrong a person by intentionally infecting his or her computer with a virus, but we do not wrong the computer itself even if we damage it irreparably and render it nonfunctional. It is possible to have duties *with regard to* some entities, such as someone's computer, without having duties *to* those entities.[8] By contrast, if we deliberately infect a person's dog with a harmful virus, we have wronged the dog's owner and also the dog. Why are persons and dogs direct moral objects and thereby distinguished from computers and houses, which are merely indirect moral objects? The answer is that direct moral objects count in their own right, are morally more than mere means to the production of benefits for others, and have basic interests,[9] whereas indirect moral objects do not. But how is the line to be drawn between what counts in its own right and what does not?

The mainstream approach has been to ask whether a being is *the kind of entity* to which moral principles or other moral categories can and should be applied and, if so, based on which *properties* of the being. In some theories, one and only one property confers moral status. For example, some say that this property is human dignity—an inexact notion that moral theory has done little to clarify. Others say that another property or perhaps several properties are needed to acquire moral status, such as sentience, rationality, or moral agency.

We argue in this chapter that the properties identified in the five most prominent theories of moral status will not, individually, resolve the main issues about moral status, but that *collectively* these theories provide a good, although untidy, framework for handling problems of moral status. We begin by looking at each of the five theories and assessing why each is attractive, yet problematic if taken to be the sole acceptable theory.

A Theory Based on Human Properties

The first theory can be called the traditional account of moral status. It holds that distinctively human properties, those of *Homo sapiens*, confer moral status. Distinctively human properties demarcate that which has moral value and delineate which beings constitute the moral community. An individual has moral status if and only if that individual is conceived by human parents—or, alternatively, if and only if it is an organism with a human genetic code. The following is a concise statement of such a position by two members of the US President's Council on Bioethics (2001–2009):

> Fertilization produces a new and complete, though immature, human organism. … A human embryo is … a whole living member of the species Homo sapiens in the earliest stage. … To deny that embryonic human beings deserve full respect, one must suppose that not every whole living human being is deserving of full respect. … [Even embryos] are quite unlike cats and dogs. … As humans they are members of a natural kind—the human species. … Since human beings are intrinsically valuable and deserving of full moral respect in virtue of what they are, it follows that they are intrinsically valuable from the point at which they come into being.[10]

Many find such a theory attractive because it unequivocally covers all human beings and demands that no human be excluded on the basis of a property such as being a fetus, having brain damage, or having a congenital anomaly. We expect a moral theory to cover everyone without making arbitrary or rigged exceptions. This theory meets that standard. The moral status of human infants, mentally disabled humans, and those with a permanent loss of consciousness (in a persistent vegetative state) is not in doubt or subject to challenge in this theory. This theory also fits well, intuitively, with the moral belief that all humans have human rights precisely because they are human.[11]

Despite its attractive features, this theory is problematic when taken as a general theory that one and only one "natural kind" deserves moral status. If we were to train nonhuman apes to converse with us and engage in moral relationships with us, as some believe has already occurred, it would be baseless and prejudicial to say that they have a lesser status merely because of a *biological* difference in species. If we were to encounter a being with properties such as intelligence, memory, and moral capacity, we would frame our moral obligations toward that being not only or even primarily by asking whether it is or is not biologically human. We would look to see if such a being has capacities of reasoning and planning, has a conception of itself as a subject of action, is able to act autonomously, is able to engage in speech, and can make moral judgments. If the individual has one or more of these properties, its moral status (at some level) is assured, whereas if it has no such properties, its moral status might be in question, depending on the precise properties it has. Accordingly, human biological properties are not necessary conditions of moral status.

Using a species criterion as the proper criterion of human properties is also not as clear and determinative as some adherents of this first theory seem to think. Consider the example of scientific research in which a monkey-human chimera is created for the purposes of stem-cell research. This research has the objective of alleviating or curing neurological diseases and injuries. It is conducted by inserting a substantial human cell contribution into a developing monkey's brain. Specifically, investigators implant human neural stem cells into a monkey's brain to see what the cells do and where they are located.[12] The question is whether functional integration of these neural cells in a nonhuman primate brain would cause a morally significant change in the mind of the engrafted animal, and, if it so, what the consequences would be for the moral status of the animal once born. Thus far, no such human-nonhuman chimera has been allowed to progress past early fetal stages, but such a chimera could be born and might be recognized as possessing a high level of moral status.

There are cells in this chimera that are distinctly human and cells that are distinctly monkey. The monkey's brain is developing under the influence of the human cells. Should it be born, it could possibly behave in humanlike ways. In theory, the larger the proportion of engrafted human cells relative to host cells, the higher the likelihood of humanlike features or responses. Such a chimera would possess a substantial human biological contribution and might have capacities for speech and moral behavior, especially if a great ape was the selected nonhuman

species.[13] Transgenic animals, that is, animals that possess and express genes from a different species, present similar issues. An example is the much-discussed Harvard oncomouse, which has only mouse cells but also has bits of human DNA and develops human skin cancers.

Related biomedical research involves the insertion of human stem cells into nonhuman animal embryos in the hope that chimeric animals containing human organs can be born and their organs transplanted into humans. These scientific studies began when stem-cell biologists successfully used injections of induced pluripotent stem cells from rats into mouse blastocysts to create mice having a rat rather than mouse pancreas.[14] This mouse-alteration research led scientists to study whether transplantable human organs might be grown in human-animal chimeras. The goal is to harvest human organs from host pig-humans in the hope that organ transplants can be made available to the hundreds of thousands of persons on waiting lists for organs around the world.[15]

The US National Institutes of Health was concerned about these studies because the injected pluripotent human cells into nonhuman embryos may have the potential to multiply and possibly to causally affect the embryo's neural development, which includes the brain, leaving "uncertainty about the effects of human cells on off-target organs and tissues in the chimeric animals, particularly in the nervous system, [which] raises ethical and animal welfare concerns."[16] We cannot decide the moral status of chimeric animals merely by the presence of possible human neural development, but it remains uncertain how best to decide these issues.[17]

There has been little opposition, other than a few concerns about human safety, to many mixtures of human and animal tissues and cells in the context of medical care (e.g., transplantation of animal parts or insertion of animal-derived genes or cells) and biomedical research (e.g., several kinds of insertion of human stem cells into animals). However, matters may become worrisome if animal-human *hybrids* are created. In 2004 the US President's Council on Bioethics found "especially acute" the ethical concerns raised by the possibility of mixing human and nonhuman gametes or blastomeres to create a hybrid. It opposed creating animal-human hybrid embryos by ex vivo fertilization of a human using animal sperm or of an animal egg using human sperm. One reason is the difficulty society would face in judging both the humanity and the moral status of such an "ambiguous hybrid entity."[18] These and other developments in research present challenges to the theory that fixed species boundaries are determinative of moral status.[19]

This first theory of moral status confronts another problem as well: The commonsense concept of *person* is, in ordinary language, functionally identical to the concept of *human being*, but there is no warrant for the assertion that only properties distinctive of the human species count toward personhood or that species membership alone determines moral status. Even if certain properties strongly correlated with membership in the human species qualify humans for moral status more readily than the members of other species, these properties are only contingently connected to being human. Such properties could be possessed by members of nonhuman species or by entities outside the sphere of natural species, such as God, chimeras, robots, and genetically manipulated species (and biological humans could, in principle, lack these properties).[20]

Julian Savulescu has proposed a way to resolve moral-status problems about the aforementioned pig-human chimeras by appeal to person theory:

> A chimera is a genetic mix. … It is not a pig with a human pancreas inserted into it—it is a human-animal chimera. … [I]t is possible that some future chimeras will develop *human or human-like brains . . . having moral relevance. . . .* If there is any doubt about the *cognitive abilities* of this new life form, we should check the chimera for its functionality. … In the absence of conclusive evidence, the default position should be that we assign them [these chimeras] *high moral status* until further research has confirmed or disproved this. …
>
> Any human-pig chimera *should, then, be assessed against the criteria of personhood*. … [A]ny such chimera should be accorded the highest moral status consistent with its likely nature.[21]

Savulescu's attention to the central place of moral status is appropriate, but it is questionable whether criteria of personhood should govern our assessments of moral status. The concept of and theory of persons is unsuited to deliver what is required unless it is convincingly argued that the concept of persons is a normative concept that can adequately resolve moral status questions. The person-theory literature is not intrinsically moral in nature, though it is also not useless in moral argument.[22] However, person theory has not proven to be the key to a satisfactory model of moral status. Moral status does not require personhood, and personhood does not clearly entail moral status, depending on what is meant by the rather imprecise notion of a "person."[23]

Some people maintain that what it means to be a person is to have some set of human biological properties; others maintain that personhood is delineated not biologically, but in terms of certain cognitive capacities, moral capacities, or both. What counts as a person expands or contracts as theorists construct their theories so that precisely the entities for which they advocate will be judged to be persons and other entities will be judged not to be persons. In one theory, human embryos are declared persons and the great apes are not, whereas in another theory the great apes are persons and human embryos are not.

The theory of moral status as grounded in properties of humanity might seem salvageable if we include both human biological properties and distinctively human psychological properties, that is, properties exhibiting distinctively human mental functions of awareness, emotion, cognition, motivation, intention, volition, and action. This broader scope, however, will not rescue the theory. If the theory is that nonhuman animals are not morally protected in a context of biomedical research because they lack psychological characteristics such as self-determination, moral motivation, language use, and moral emotions, then consistency in theory requires stating that humans who lack these characteristics likewise do not qualify for moral protections for the same reason. For any human psychological property we select, some human beings will lack this characteristic (or at least lack it to the relevant degree); and frequently some nonhuman animal will possess this characteristic. Primates, for example, often possess humanlike properties that some humans lack, such as a specific form of intellectual quickness, the capacity to feel pain, and the ability to enter into meaningful social relationships. Accordingly, this first theory based on human properties does not by itself qualify as a comprehensive account of moral status.

Nonetheless, it would be morally perilous to give up the idea that properties of humanity form a basis of moral status. This position is entrenched in morality and provides the foundation of the claim that all humans have human rights. Accordingly, the proposition that *some* set of distinctive human properties is *a sufficient, but not necessary, condition of moral status* is an attractive and we think acceptable position.[24] However, we leave it an open question precisely which set of properties counts, and we acknowledge that argument is needed to show that some properties count whereas others do not. We also acknowledge that it could turn out that the properties we regard as the most critical human properties are not distinctively human at all.

The acceptance of a criterion of human properties as supplying a sufficient condition of moral status does not rule out the possibility that properties other than distinctively human ones also constitute sufficient conditions of moral status. To test this hypothesis, we turn to consideration of the other four theories.

A Theory Based on Cognitive Properties

A second theory of moral status moves beyond biological criteria and species membership to cognitive properties that are often associated with the properties of being a person. "Cognition" refers to processes of awareness such as perception, memory, understanding, and thinking. This theory does not assume that only humans have such properties, although the starting model for these properties is usually the competent human adult. The theory is centrally that individuals have moral status because they are able to reflect on their lives through their cognitive capacities and are self-determined by their beliefs in ways that incompetent humans and many nonhuman animals are not.

Properties found in theories of this second type include (1) self-consciousness (consciousness of oneself as existing over time, with a past and future); (2) freedom to act and the capacity to engage in purposeful actions; (3) ability to give and to appreciate reasons for acting; (4) capacity for beliefs, desires, and thoughts; (5) capacity

to communicate with other persons using a language; and (6) rationality and higher order volition.[25] The goal of theories of this type is to identify a set of cognitive properties possessed by all and only beings having moral status. We here set aside disputes internal to these theories about precisely which cognitive properties are jointly necessary and/or sufficient for personhood, and therefore for moral status. To investigate the problems with this general type of theory, it does not matter for present purposes whether only one or more than one of these properties must be satisfied.

The model of an autonomous human being, or person, is conceived in many of these theories in terms of cognitive properties such as those listed in the previous paragraph. The theory that these properties form the foundation of moral status acknowledges that if a nonhuman animal, a hybrid human, or a brain-damaged human is in all relevant respects like a cognitively capable human being, then it has a similar (presumably identical) moral status. A corollary is that if one is *not* in the relevant respects similar to a cognitively competent human being, one's moral status is correspondingly reduced or vacated.

As the number or level of the required cognitive abilities is increased, a reduction will occur in the number of individuals who satisfy the theory's conditions, and therefore fewer individuals will qualify for moral status or at least for elevated moral status. For example, if all six of the previously listed criteria must be satisfied, many humans would be excluded from elevated moral status. Likewise, if the quality or level of the required cognitive skills is reduced, the number of individuals who qualify for protection under the theory will presumably increase. For example, if only understanding and intentional action at a basic level were required, some nonhuman animals would qualify.

A worrisome feature of this theory is that infants, the senile elderly, persons with a severe mental disability, and others who are generally regarded as having a secure moral status will lack the cognitive capacities required to attain moral status. Most nonhuman animals may also lack these cognitive capacities. The level of cognitive abilities required also may vary from one theory to the next theory. In explicating a Kantian position, Christine Korsgaard writes, "Human beings are distinguished from animals by the fact that practical reason rather than instinct is the determinant of our actions."[26] If this criterion of practical reason were the sole criterion of moral status, then biological "humans" who lack practical rationality would be mere animals (and not even truly human beings).

An objection to this theory, often directed against theories predicated primarily on human dignity or autonomy, is "the argument from marginal cases." This argument maintains that every major cognitive criterion of moral status (intelligence, agency, self-consciousness, etc.) excludes some humans, including young children and humans with serious brain damage. These "marginal" cases of cognitive human capacities can be at the same level of cognitive (and other) capacities as some animals, and therefore to exclude these animals is also to exclude comparably situated humans. If animals can be justifiably treated as mere means to human ends, then comparable "marginal" cases of human capacity can also be justifiably treated as mere means to human ends—for example, by becoming research subjects.[27] This position precludes a high level of moral status for many weak, vulnerable, and incapacitated humans.

This theory therefore does not function, as the first theory does, to ensure that vulnerable human beings will be morally protected. The more vulnerable individuals are by virtue of cognitive deficiency, the weaker are their claims for moral protection. The fact that members of the human species typically exhibit higher levels of cognitive capacities than members of other species does not alleviate this problem. Under this theory, a nonhuman animal in principle can overtake a human in moral status once the human loses a measure of mental abilities after a cataclysmic event or a decline of capacity. For example, once a primate training in a language laboratory exceeds a deteriorating Alzheimer's patient on the relevant scale of cognitive capacities, the primate would attain a higher moral status in this type of theory.[28]

Writers in both science and biomedical ethics often assume that nonhuman animals lack the relevant cognitive abilities, including self-consciousness (even basic consciousness), autonomy, or rationality, and are therefore not elevated in status by this theory.[29] However, this premise is more assumed than demonstrated. Much has been demonstrated about cognition in animal minds by ethologists who investigate animal cognition and mental

properties using evolutionary and comparative studies as well as naturalistic and laboratory techniques of observation and experimentation.[30] Comparative studies of the brain show many relevant similarities between the human species and various other species. In behavioral studies, some great apes appear to make self-references or at least to show self-awareness or self-recognition, and many animals learn from the past and use their knowledge to forge intentional plans of action for hunting, stocking reserve foods, and constructing dwellings.[31] In play and social life, many animals understand assigned functions and either follow designated roles or decide for themselves what roles to play.[32] Moreover, many animals seem to understand and intend in ways that some incapacitated humans cannot. These are all *cognitively* significant properties, and therefore, in this second theory, they are *morally* significant properties that award a more elevated moral status to nonhuman animals with the relevant properties than to humans who lack them.

Defenders of this second type of theory need to address how to establish the relevance and importance of the connection asserted between cognitive properties and moral protections. Why do *cognitive* properties of individuals determine anything at all about their *moral* status? We are not asserting that a theory of moral status cannot be based on nonmoral properties. It can, but such a theory of moral status must make a connection between its preferred nonmoral properties and the claim that they confer moral status. Defenders need to explain why the absence of this property (e.g., self-consciousness) makes a critical moral difference and precisely what that difference is. If a human fetus or an individual with advanced dementia lacks certain cognitive properties, it does not follow, without supporting argument, that they lack moral status and associated moral protections.

To conclude this section, this second theory, like the first, fails to establish that cognitive capacity is a *necessary condition* of moral status. However, the theory arguably does succeed in showing that some set of cognitive capacities is a *sufficient condition* of moral status. Cognitive capacities such as reasoned choice occupy a central place in what we respect in an individual when we invoke moral principles such as "respect for autonomy." The main problem with this second theory is not that it invokes these properties, but that it considers *only* cognitive properties and neglects other potentially relevant properties, notably properties on the basis of which individuals can suffer and enjoy well-being. We will see below in examining the fourth theory of moral status that certain noncognitive properties are also sufficient for moral status.

A Theory Based on Moral Agency

In a third type of theory, moral status derives from the capacity to act as a moral agent. The category of *moral agency* is subject to different interpretations, but, fundamentally, an individual is a moral agent if two conditions are satisfied: (1) the individual is capable of making moral judgments about the rightness and wrongness of actions, and (2) the individual has motives that can be judged morally. These are moral-capacity criteria, not conditions of morally correct action or character. An individual could make immoral judgments and have immoral motives and still be a moral agent.[33]

Several theories fall under this general type, some with more stringent conditions of moral agency than the two just listed. Historically, Immanuel Kant advanced what has become the most influential theory of moral agency. He concentrated on moral worth, autonomy, and dignity, but some of his formulations suggest that he is also proposing conditions of moral status. For example, moral autonomy of the will is central to his theory. It occurs if and only if one knowingly governs oneself in accordance with universally valid moral principles. This governance gives an individual "an intrinsic worth, i.e., dignity," and "hence autonomy is the ground of the dignity of human nature and of every rational creature."[34]

Kant and many after him have suggested that capacity for moral agency gives an individual a moral respect and dignity not possessed by individuals incapable of moral agency—human or nonhuman. This account has a clearly attractive feature: Being a moral agent is indisputably a *sufficient* condition of moral status. Moral agents are the paradigmatic bearers of moral status. They know that we can condemn their motives and actions, blame them for irresponsible actions, and punish them for immoral behavior.[35]

Accordingly, like the first two theories, this third theory supplies a sufficient condition of moral status, and, like the first two, it fails to identify a *necessary* condition of moral status. If being a moral agent (or being morally autonomous) were a necessary condition of moral status, then many humans to whom moral protections are extended would be stripped of their moral status, as would most and perhaps all nonhuman animals. Many psychopaths, patients with severe brain damage, patients with advanced dementia, and animal subjects in research would lack moral status in this theory. Yet individuals in these classes deserve to have their interests attended to by many parties, including institutions of medical care. The reason for such protections cannot be a capacity of moral agency, because these individuals have none.

Interpreting the theory of moral agency as a necessary condition of moral status is strongly counterintuitive. A morally appropriate response to vulnerable parties such as young children, the severely intellectually disabled, patients with senile dementia, and vulnerable research animals is that they deserve *special* protection, not that they merit no protection. Whether these individuals are moral agents is not the primary consideration in assessing their moral status.

Accordingly, this third theory provides a sufficient condition of moral status but not a necessary one. We have already seen that there are other ways to acquire moral status, and we will now argue that a fourth theory lends additional support to this conclusion.

A Theory Based on Sentience

Humans as well as nonhuman animals have properties that are neither *cognitive* nor *moral* properties, yet count toward moral status. These properties include a range of emotional and affective responses, the single most important being *sentience*—that is, the capacity for consciousness understood as experience in the form of feelings. Specifically, sentience is the capacity for sensations, feelings, or other experiences that are agreeable or disagreeable. Because sentient animals have a subjective quality of life, they have an experiential welfare and therefore welfare interests.[36]

A central line of moral argument in this fourth theory is the following: Pain is an evil, pleasure a good. To cause pain to any entity is to harm it. Many beings can experience pain and suffering, which are bad in themselves and even worse when experienced over an extended period of time.[37] To harm these individuals is to *wrong* them, and such harm-causing actions are morally prohibited unless one has moral reasons sufficient to justify them.

Proponents of this fourth theory appropriately claim that having the capacity of sentience is a sufficient condition of moral status.[38] The properties of being able to experience pain and suffering are almost certainly sufficient to confer some measure of moral status. One of the main objectives of morality is to minimize pain and suffering and to prevent or limit indifference and antipathy toward those who are experiencing pain and suffering. We need look no further than ourselves to appreciate this point: Pain is an evil to each of us, and the intentional infliction of pain is a moral-bearing action from the perspective of anyone so afflicted. What matters, with respect to pain, is not species membership or the complexity of intellectual or moral capacities. It's the pain. From this perspective, all entities that can experience pain and suffering have some level of moral status.

This theory has broad scope. It reaches to vulnerable human populations and to many animals used in biomedical research. We study animals in biomedical research because of their similarities with humans. The reason to use animals in research is that they are so similar to humans, and the reason not to use animals in research is that they are so similar to humans in their experience of pain and suffering. Notably in the case of primates, their lives are damaged and their suffering often resembles human suffering because they are similar to us physically, cognitively, and emotionally.

Precisely *who* or *what* is covered by this conclusion, and *when*, is disputed, especially in the large literatures on animal research, human fetal research, and abortion. If sentience alone confers moral status, a human fetus acquires moral status no earlier and no later than the point of sentience. Growth to sentience in the sense of a biological process is gradual over time, but the acquisition of sentience—or the first onset of sentience—is, in this fourth theory, the point at which moral status is obtained. Some writers argue that development of a

functioning central nervous system and brain is the proper point of moral status for the human fetus, because it is the initial biological condition of sentience.[39] This approach does not protect human blastocysts or embryos and has proved to be an uncertain basis on which to build arguments allowing or disallowing abortion, because there is disagreement about when the brain has developed sufficiently for sentience. However, in this theory a fetus acquires moral status at some point after several weeks of development, and thus abortions at that point and later are (prima facie) impermissible.[40] We are not, in making these observations, presenting objections to sentience theory or to any version of it. We are noting only that these problems need to be addressed in a comprehensive theory of moral status that emphasizes sentience.

Defenders of a sentience theory often quote Jeremy Bentham's famous statement: "The question is not, Can they *reason*? nor, Can they *talk*? but, Can they *suffer*?"[41] Advocates emphasize that moral claims on behalf of any individual, human or nonhuman, may have nothing to do with intelligence, capacity for moral judgment, self-consciousness, rationality, personality, or any other such fact about the individual. The bottom line is that sentience is a *sufficient* condition of moral status independent of these other properties of individuals.

The theory that sentience is a sufficient condition of moral status makes more modest claims than the theory that sentience is a necessary and sufficient condition and thus the only criterion of moral status. The latter theory is embraced by a few philosophers who hold that properties and capacities other than sentience, such as human biological life and cognitive and moral capacities, are not defensible bases of moral status.[42] Nonsentient beings, such as computers, robots, and plants (and also nonsentient animals), lack the stuff of moral status precisely because they have no capacity for pain and suffering; all other beings deserve moral consideration because they are sentient.

This very strong version of the fourth theory is problematic. The main problem arises from the claim that an individual lacking the capacity for sentience lacks moral status. On the human side, this theory disallows moral status for early-stage fetuses as well as for all humans who have irreversibly lost the capacity for sentience, such as patients with severe brain damage. It is not satisfactory to assert that absence of sentience entails absence of moral status. Proponents of the sentience theory might seek to defend it in several ways, probably by accepting another criterion of moral status in addition to that of sentience. This maneuver would give up the claim that sentience is a necessary and sufficient condition of moral status, which would be to abandon robust theories of the fourth type.

Another problem with strong versions of the fourth theory is their *impracticability*. We could not hope to implement these versions in our treatment of all species whose members are capable of sentience, and we could not do so without presenting grave danger to human beings. Virtually no one defends the view that we cannot have public health policies that vigorously control for pests and pestilence by extermination. The most plausible argument by a sentience theorist who holds the view that sentience is sufficient for moral status is that the theory grants only *some level* of moral status to sentient beings.

The most defensible theory of this fourth type holds (1) that not all sentient creatures have the same level of sentience and (2) that, even among creatures with the same level of sentience, sentience may not have the same significance because of its interaction with other properties. A few writers believe that there is a gradation of richness or quality of life, depending on level of consciousness, social relationships, ability to derive pleasure, creativity, and the like. A continuum of moral status scaled from the autonomous adult human down through the lowest levels of sentience can, in this way, be layered into sentience theory. Even if many sentient animals have moral status, it does not follow that humans should be treated no differently than other animals including the great apes. There may be many good reasons for forms of differential treatment.

In one such theory a human life with the capacity for richness of consciousness has a higher moral status and value than even a richly flourishing animal life such as that of a dog or a bonobo. This judgment has nothing to do with species membership, but rather with "the fact that [rich, conscious] human life is more valuable than animal life" by virtue of capacities such as genuine autonomy. In this theory human life is valuable and has moral status only under certain conditions of quality of life. Human life, therefore, can lose some of its value and moral status by degrees as conditions of welfare and richness of experience decrease.[43] All such theories have

problems that need resolution because the moral status of a life and its protections decline by degrees as conditions of welfare and richness diminish. When loss of capacity occurs, for example, humans and nonhumans alike will have a reduced moral status, and the most vulnerable individuals will become more susceptible to abuse or exploitation because of their reduced moral status. No theory that supports this conclusion in general is morally acceptable.

In light of the several problems surrounding the theory that sentience is both a necessary and sufficient condition of moral status, we conclude that this fourth theory—like the first three theories—provides a sufficient, but not a necessary, condition of some level of moral status. This theory needs supplementation by the other theories previously discussed to provide a comprehensive account of moral status. Sentience theory can be used to determine which beings have moral status, whereas other theories could help determine the degree of moral status. Unless augmented, this fourth theory does not determine the precise level of moral status or the proper scope of moral protections.

A Theory Based on Relationships

A fifth and final theory is based on relational properties. This theory holds that relationships between parties confer moral status, primarily when relationships establish roles and obligations. An example is the patient-physician relationship, which is a relationship of medical need and provision of care. Once this relationship is initiated, the patient gains a right to care from this particular physician lacked by persons who are not the physician's patient. The patient does not have this status independent of an established relationship, and the physician does not have the same obligations to those outside the relationship.

Other examples are found in relationships that do *not* involve a formal understanding between the parties, such as bonds with persons with whom we work closely and relationships that involve no mutual understanding between the parties, such as human initiatives that establish relations with laboratory animals and thereby change what is owed to these animals. A much-discussed example is the relationship between human personnel in laboratories and animal subjects who are thoroughly dependent on their caretakers. Here the caretaker role generates obligations on investigators and other responsible parties.

This fifth theory tries to capture the conditions under which many relationships in research and practice, especially those involving social interaction and reciprocity, are stronger and more influential than relationships with strangers and outsiders. One version of this theory depicts the relevant relationships as developing in diverse ways over time. Alzheimer's patients and experimental animals, for example, have a history in which the human moral community has assessed the importance of its relationship to these individuals. In each case we owe protection and care to those with whom we have established these relationships, and when they are vulnerable to harm we have special obligations to protect and care for them because of these relationships.[44]

In some versions of this theory, the human fetus and the newborn baby are examples of those who gradually come to have a significant moral status through special social relationships. Here is one such account of the moral status of the human fetus:

> The social role in question develops over time, beginning prior to birth. … A matrix of social interactions between fetus and others is usually present well before parturition. Factors contributing to this social role include the psychological attachment of parents to the fetus, as well as advances in obstetric technology that permit monitoring of the health status of the fetus. … The less the degree to which the fetus can be said to be part of a social matrix, the weaker the argument for regarding her/him as having the same moral status as persons. Near the borderline of viability, … the fetus might be regarded as part of a social network to a lesser degree than at term. If so, the degree of weight that should be given to the fetus's interests varies, being stronger at term but relatively weaker when viability is questionable.[45]

Despite its attractions, this fifth theory cannot do more than account for how moral status and associated protections are *sometimes* established. If this theory were taken as the sole basis of moral status, only social

bonds and special relationships would determine moral status. Critical rights such as the right to life and the right not to be confined have no force in such a theory unless rights are conferred in a context of relationships. The theory is unsustainable as an account of moral status if it rejects, neglects, or omits the insights in the previous four theories, which recognize moral status on the basis of qualities (cognition, sentience, etc.) that can be acknowledged independently of relationships. For example, in the fourth theory, the property of sentience is status conferring. When we wrongfully harm a human research subject or a human population through environmental pollution, it is incorrect to say that the harming is wrong merely because we have an established laboratory, clinical, or social relationship with either particular individuals or populations. We behave wrongly because we cause gratuitous and unnecessary risk, pain, or suffering, which would be so whether or not an established relationship exists.

The problem of moral status is fundamentally about which beings have moral status, and this fifth theory does not directly address this problem. It rather focuses on the basis on which beings sometimes gain or lose specific moral rights or generate or discontinue specific moral obligations. Accordingly, this fifth theory does not supply a necessary condition of moral status, and, in contrast to the other theories we have examined, it also does not clearly provide a sufficient condition of moral status in many cases of important relationships.[46] Many loving and caring relationships, with various kinds of beings, do not confer moral status on those beings. No matter how much we love our children's closest friends or a neighbor's pet, they do not gain moral status by virtue of our relationship to them. Nor does the lack of such a relationship indicate a lack of moral status. An individual still may gain status under criteria drawn from one of the four previous theories (humanity, cognition, moral agency, and sentience). This approach is the best way to maximally preserve claims of moral status for individuals no longer capable of having significant interpersonal relationships. They will not lose all moral status merely because relationships have been lost.

In sum, the fifth theory's primary contribution is to show that certain relationships account for how many individuals acquire or lose some moral entitlements and others engender or discontinue obligations. In this way, the theory helps account for different degrees of moral status, as discussed in the section below on "Degrees of Moral Status**.**"

FROM THEORIES TO PRACTICAL GUIDELINES

Each of the five theories examined thus far has acceptable and attractive elements. However, each theory risks making the mistake of isolating a singular property or type of property—biological species, cognitive capacity, moral agency, sentience, or special relationships—as the sole or at least the primary criterion of moral status. Each theory proposes using its preferred property for including certain individuals (those having the property) and excluding others (those lacking the property). Each theory thereby becomes too narrow to be a general theory of moral status unless it accepts some criteria in one or more of the other four theories.

From ancient Hellenic times to the present, we have witnessed different motives and theories at work when groups of people (e.g., slaves and women) have been denied a certain social standing because they lack some highly valued property that would secure them full moral status. Over time, views about the moral acceptability of these presumed criteria have changed and have altered beliefs about the moral status of members of these groups. For example, women and minority groups denied equal moral status later received, in many societies, the equal status that ought never to have been denied. The worry still today is that some groups, especially vulnerable groups including some patients and research subjects, still face a discriminatory social situation: They fail to satisfy criteria of moral status because the dominant criteria have been tailored specifically so that they do not qualify for full—or perhaps even partial—moral status. Discussion in biomedical ethics has focused principally on whether the following are vulnerable groups of this description: human embryos, human fetuses, anencephalic children, human research subjects, animal research subjects, and individuals affected by unresponsive wakefulness syndrome (or persistent vegetative state).[47]

The primary norms in each theory—which we hereafter refer to as *criteria* of moral status (rather than *theories* or *conditions* of moral status)—work well for some problems and circumstances in which decisions must be

made, but not well for other problems and circumstances.

Appropriation of the Best Criteria from the Five Theories

Ideally, we can adopt the best from each of the five theories and meld these elements into a multicriterial, coherent account of moral status.[48] This strategy will help accommodate the diversity of views about moral status, will allow a balancing of the interests of different stakeholders such as the interests of scientists in new knowledge and the interests of research subjects, and will help avoid intractable clashes of rights, such as conflicts between the rights of scientists to engage in research and the rights of human embryos. We hereafter assume that, in principle, the ideal of a coherent, multicriterial account of moral status can be satisfied; but a unified and comprehensive account of moral status is a demanding and ambitious project that we make no claim to have undertaken in the present chapter.

Degrees of Moral Status

In many accounts of moral status, not all individuals enjoying moral status have it categorically, without qualification, or fully. In some theories, competent, adult humans have a broader array of rights than other beings, especially rights of self-determination and liberty, because of their capacities of autonomy and moral agency. Despite the now common view that many species of animals involved in research have some level of moral status, it is rare to find a theory of moral status that assigns all animals in research the same degree of moral status as human persons.[49] Even defenders of animal rights generally acknowledge that it is worse to exterminate a person than to exterminate a rat. Another common view is that frozen human embryos do not have the same moral status as human persons. But are these claims about higher and lower moral status defensible? Does a defensible theory recognize degrees of moral status?

We start toward an answer by examining a groundbreaking case in public policy that relies on the idea of degrees of moral status. This case derives from the history of debate and legislation about human embryo research in the United Kingdom. The morally contentious issues surrounding this research were first considered by the Committee of Inquiry into Human Fertilisation and Embryology (the Warnock Committee, 1984)[50] and later debated in Parliament during passage of the Human Fertilisation and Embryology Act of 1990. Regulations in 2001 set regulatory policy governing the use of embryos in research. These regulations were indebted to a 2000 report by the Chief Medical Officer's Expert Group.[51] According to this report, British policy affirms the following moral principles as the moral basis of law and regulation regarding the use of embryos in stem-cell research:

> The 1990 Act reflects the majority conclusion of the Warnock Committee. The use of embryos in research in the UK is currently based on the [following] principles expressed in their Report:
>
> - The embryo of the human species has a special status but not the same status as a living child or adult.
> - The human embryo is entitled to a measure of respect beyond that accorded to an embryo of other species.
> - Such respect is not absolute and may be weighed against the benefits arising from proposed research.
> - The embryo of the human species should be afforded some protection in law. …
>
> The Expert Group accepted the 'balancing' approach which commended itself to the majority of the Warnock Committee. On this basis, extending the permitted research uses of embryos appears not to raise new issues of principle.[52]

This position is a somewhat vague, but common—and, in this case, a highly influential—expression of an account of degrees and levels of moral status and concomitant protections.

The five theories we have addressed can each be interpreted in terms of degrees. For example, in the fourth theory, based on sentience, moral status is arguably proportional to degree of sentience and perhaps to the quality and richness of sentient life. Similarly, in the fifth theory, based on relationships, moral status is expressible in terms of degrees of relationship: Relationships come in different degrees of closeness, and relations of dependence can be far more significant in some cases than in other cases.

Arguably, all morally relevant properties in each of these theories are degreed. Capacity for language use, sentience, moral agency, rationality, autonomous decision making, and self-consciousness all come in degrees and may not be limited to human beings.[53] From this perspective, there are higher and lower levels of moral status, and we can conceive a continuum running from full moral status to no moral status.

But is an account of degrees of moral status superior to an all-or-nothing account of moral status?[54] The notion of a lesser moral status (including the notion of being subhuman or inhuman) has been troublesome throughout history, and its remnants linger in many cultural practices. Is it, then, best to deny or to affirm that there are degrees of moral status?

These problems of degrees of moral status should not obscure the fact that all beings with moral status, even those unambiguously below full moral status, still have *some* significant moral status. Disagreement is inevitable regarding whether the concept of degrees is suitable for the analysis of all properties that confer moral status. For example, disagreement appears in the writings of those having firm commitments to the first theory, based on properties of humanity. One controversial case involves the potential of a human fetus to become a sentient, cognitively aware, moral agent. In some theories this potential is not expressible by degrees because full potential is present from the start of an individual's life; a human fetus therefore has full moral status at its origins and throughout its existence. In other theories human fetuses have a lower degree of moral status because they are only potential persons, not yet actual persons.

In one type of theory, the moral status of human zygotes, embryos, and fetuses increases gradually during gestation.[55] This theory can be developed to make potentiality itself a matter of degree (degree of potentiality). For example, brain defects in a fetus or infant can affect the potential for cognitive and moral awareness and also for the relationships that can be formed with others. This theory can also be expressed in terms of different sets of rights—for instance, pregnant women may have more rights than their fetuses as well as a higher level of moral status than their fetuses—at least at some stages of fetal development.

A practically oriented theory of moral status will need to determine with precision what an individual's or a group's status is, not merely that the individual or group has some form of status. A comprehensive theory will explain whether and, if so, how the rank will change as properties that contribute to status are progressively gained or lost. We ought not to be optimistic that such a theory can be developed to cover all problems of moral status, but we can hope to achieve a better theory than has thus far been available.

The Connection between Moral Norms and Moral Status

We have distinguished questions about moral status from the questions about the moral norms addressed in Chapter 1. We will now further develop this distinction. Criteria of moral status *are* moral norms in the generic sense of "moral norm." A moral norm in the generic sense is a (prima facie) standard that has the authority to judge or direct human belief, reasoning, or behavior. Norms guide, require, or commend. Failure to follow a norm warrants censure, criticism, disapproval, or some other negative appraisal. Criteria of moral status satisfy this description. Although not the same type of norm as principles and rules, these criteria are normative standards.

Criteria of moral status also can be understood in terms of the discussions in Chapter 1 of moral conflict, moral dilemmas, prima facie norms, and the specification and balancing of norms. Criteria of moral status can and often do come into conflict. For example, the criterion of sentience (drawn from theory 4) and the criterion of human species membership (drawn from theory 1) come into conflict in some attempts to determine the moral status of the early-stage human fetus. The sentience criterion expressed in theory 4 suggests that the fetus gains

status only at the point of sentience, whereas the criterion of human properties (in theory 1) suggests that moral status accrues at human biological inception.

Guidelines Governing Moral Status: Putting Specification to Work

Conflicts of theory and interpretation can and should be addressed using the account of specification delineated in Chapter 1. Norms are specified by narrowing their scope, which allows us to create what we will call *guidelines* governing moral status. Others might call them rules instead of guidelines, but in our framework rules specify principles whereas guidelines specify criteria of moral status. The goal is to extract content from the criteria found in one or more of the five theories to show how that content can be shaped into increasingly practical guidelines. We will state these guidelines using the language of a "level of moral status."

The concept of a level should be interpreted in terms of degrees of moral status. This approach provides for a continuum of moral status, running from a narrow range of moral protections to a broad range of moral protections. For example, infants, the mentally handicapped, and many persons who are cognitively incompetent have some level of moral status, but they do not have the same level of moral status as autonomous persons. For instance, those who lack substantial cognitive and autonomy capacities will not have various decision-making rights such as the right to give an informed consent that are enjoyed by those who are substantially autonomous, but they will still have rights to life and to health care. To say that they have a lower moral status is not to demean or degrade them. It is to recognize that they do not have the same entitlements that others have. But their vulnerabilities also may confer entitlements on them that others do not have such as various entitlements to medical care and special education.

To show how norms can be made progressively practical, we will now treat illustrative specifications that qualify as guidelines. We are not recommending the five guidelines below. Our goal is merely to clarify the nature, basis, and moral significance of these guidelines and to show how they are formed using the method of specification.

Consider first a circumstance in which the criterion "All living human beings have some level of moral status" comes into conflict with the criterion "All sentient beings have some level of moral status." We start with two possible specifications (guidelines 1 and 2 below) that engage the criteria put forward in theories 1 (the criterion of human life) and 4 (the criterion of sentience):

> **Guideline 1.** All human beings who are sentient or have the biological potential for sentience have some level of moral status; all human beings who are not sentient and have no biological potential for sentience have no moral status.

This specification allows for additional specification applicable to particular groups such as brain-dead individuals, anencephalic individuals (those without a cerebrum and cerebellum, which are essential to significant levels of thinking and behavior), and individuals who have sufficient brain damage that they are not sentient and have no potential for sentience. Guideline 1 says that individuals in such groups have no moral status. By contrast, the guideline assigns some level of moral status to all healthy human embryos and fetuses when they are either sentient or have the potential to be sentient. Guideline 1 cannot be used to support human embryonic stem-cell research or abortions and so might not support the transplantation of human fetal stem cells into a Parkinson's patient. Guideline 1 stands opposed to these practices, though it too can be further specified.

A different, and obviously *competitive*, guideline that is achieved through specification is this:

> **Guideline 2.** All human beings who are sentient have some level of moral status; all human beings who are not sentient, including those with only a potential for sentience, have no moral status.

This second guideline has profoundly important moral implications for whether embryos and early-stage fetuses have moral status and therefore implications for moral debates about human embryonic stem-cell research and early-stage abortions. It states that although life prior to sentience is morally unprotected, the fetus is protected

against abortion and research interventions once it becomes sentient.[56] Unlike guideline 1, guideline 2 would allow the transplantation (after appropriate research) of human fetal stem cells into a Parkinson's patient.

Clarifying the exact implications of this second guideline would require further specification(s). In the case of abortion in particular, even when a fetus is sentient its continued existence could threaten the life or health of the pregnant woman. On one possible line of further specification, sentient fetuses possess the *same* rights possessed by all sentient human beings, and an abortion is a maleficent act as objectionable as the killing of an innocent person. On a different line of specification, sentient fetuses have a diminished set of rights if their presence threatens the life of a pregnant woman. In the abstract form here presented, guideline 2 is only a first step in grappling with problems governing several classes of individuals.

A third possible guideline reached by specification appeals both to theory 4 (sentience) and to theory 2 (cognitive capacity):

> **Guideline 3.** All sentient beings have some level of moral status; the level is elevated in accordance with the level of sentience and the level of cognitive complexity.

According to this guideline, the more sentient the individual and the richer the cognitive or mental life of the individual, the higher the individual's level of moral status. The capacities of creatures for an array of valuable experiences vary. As a result, not all lives are lived at the same high level of perception, cognition, appreciation, esthetic experience, and the like. The issue is not whether a life has value; it is about different levels of value because of differences in sentience and the quality of mental life. This guideline is a first step toward working out the common intuition in research involving animals that great apes deserve stronger protections than pigs, which deserve more protection than rats, and so forth. However, this guideline might not turn out to support many common intuitions about the mental capacities of species; for example, pigs could turn out to have a richer mental life than dogs or baboons and therefore a higher moral status than members of these species.[57]

Depending on how this guideline is further specified, it might or might not support use of a ready-to-transplant pig heart valve into a human heart. The level of the pig's capacities of sentience and cognition might make a critical moral difference in whether the valve can be harvested from pigs in the first place. Under this guideline, questions of the comparative value of the human life saved and the sacrificed pig's life can only be decided by inquiry into the levels of their sentience and cognition.

Consider now a fourth guideline, this one a specification of the criterion of moral agency (theory 3) in conflict with the criterion of human-species properties (theory 1):

> **Guideline 4.** All human beings capable of moral agency have equal basic rights; all sentient human beings and nonhuman animals not capable of moral agency have a diminished set of rights.

This guideline sharply elevates the status of moral agents while giving a lesser status to all other sentient creatures. Defense of this guideline would likely require an account of equal basic rights and of which rights are held and not held by those incapable of moral agency (a subject partially treated in Chapter 4).

This guideline is, from one perspective, obviously correct and noncontroversial: Competent individuals capable of moral agency have a set of rights—for example, decision-making rights—not held by individuals who are not capable of moral agency, whether the latter are human or nonhuman. Far more controversial and difficult to handle by specification is the underlying premise that human individuals who lack capacity for moral agency thereby have a reduced moral status. Proponents of theory 1 presumably would altogether reject this premise in their specifications. Categorization of reduced moral status could affect many decisions in bioethics such as how to rank order who has primacy in the order of who receives organ transplants (under conditions of scarcity of organs). A lingering question would be whether individuals with no capacity for moral agency should be accorded a reduced moral status that ranks them sufficiently low that they are not competitive for transplantation.

Consider, as a final example, a possible guideline that engages the demands of the fifth theory (of status through relationships) and the fourth theory (of sentience). This specification brings the two criteria to bear on the circumstance of laboratory animals. The following formulation assumes the moral proposition that the "communal relationship" between persons in charge of a laboratory and the animals in it is morally significant:

> **Guideline 5.** All sentient laboratory animals have a level of moral status that affords them some protections against being caused pain, distress, or suffering; as the likelihood or the magnitude of potential pain, distress, or suffering increases, the level of moral status increases and protections must be increased accordingly.

This guideline is the first step in making precise the idea that laboratory animals who benefit human communities gain a higher moral status than would the same animal having only sentience. Laboratory rats, for example, gain more status than rats living in the woods or in the attics of hospitals. Human initiatives that establish relations with animals change what is owed to them, and they thereby acquire a higher status than do wild animals of the same species. The main conditions of interest are the vulnerability and dependence engendered in animals when humans establish relations with them in laboratories. The more vulnerable research makes the animals to pain and suffering, the more obligations of animal care and protection increase.

This guideline has sometimes been expressed in terms of human stewardship over the animals—that is, the careful and responsible oversight and protection of the conditions of an animal entrusted to one's care. However, a better model—because of its closeness to moral status criteria—is grounded in obligations of reciprocity and nonmaleficence: Animal research subjects gain a higher moral status because of the use made of their bodies and the harm or risk of harm in the research.

These five guidelines might be presented in such abstract and indeterminate formulations that they will seem doubtfully practicable. If their abstractness cannot be further reduced, this outcome would be unfortunate because practicability is an important standard for evaluation of all accounts in practical ethics. In principle guidelines can be progressively specified to the point of practicability, just as moral principles can (as demonstrated in Chapter 1). In addition, constrained balancing (also analyzed in Chapter 1) will often have a role in determining justifiable courses of action.

THE MORAL SIGNIFICANCE OF MORAL STATUS

Some writers challenge the need for the category of moral status. They argue that moral theory can and should move directly to guidance about how individuals ought to be treated or to which moral virtues should be enacted. Some philosophers argue that moral status accounts of the sort examined thus far offer a superficially attractive but overly simplistic picture of how we "expand the circle of our concern" beyond autonomous adult humans to human fetuses, brain-damaged humans, laboratory animals, and the like. They argue that such theories blind us to the range of features that are morally relevant in decision making. If a creature has a property such as sentience, this fact does not tell us how we should treat or otherwise respond to members of the class of sentient beings; nor does it give us an account of moral priorities. Accordingly, we do not need the concept and theory of moral status and would be better off without it.[58]

This account proposes that we attend to various morally relevant features of situations that give us reasons for acting or abstaining from acting in regard to others that no theory of moral status is well equipped to address. For example, we often make distinctions that lead us to justifiably give preferential treatment to either individuals or classes of individuals, such as preferences to our children, our friends, our companion animals, and the like. We have to sort through which preferences are justifiable and which not, but no general theory of moral status suitably directs us in this task.

These cautions appropriately warn us about the limits of theories of moral status, but moral status remains a matter of paramount moral importance and should be carefully analyzed, not ignored or downplayed. We take a similar view about basic human rights in Chapter 9. It would be a catastrophic moral loss if we could not be guided by basic norms of moral status and basic rights. Practices of slavery as well as abuses of human research

subjects have thrived historically in part because of defective criteria of moral status and inattention to basic rights connected to moral status. In too many places in recent decades, some children who were institutionalized as "mentally infirm," some elderly patients in chronic disease hospitals, and some racial groups were treated as if they had little or no moral status by some of the finest centers of biomedical research in the world and by the sponsors of such research.[59] It is easy to forget how recognition of moral status can generate interest in and support acknowledgment of vital moral protections.[60]

VULNERABLE POPULATIONS AND VULNERABLE INDIVIDUALS

Concern about moral status has often arisen from the need to protect vulnerable populations. Rules requiring additional protections for certain populations are a foundation stone of both clinical ethics and research ethics. These protections arose historically from concerns about exploitation and the inability of the members of some groups to consent to or to refuse an intervention.[61] Vulnerable persons in biomedical contexts are sometimes incapable of protecting their interests because of sickness, debilitation, mental illness, immaturity, cognitive impairment, and the like. They may be socioeconomically impoverished, which adds to the potential for harmful outcomes. Populations such as homeless families, political refugees, and illegal aliens can also in some circumstances be considered vulnerable.

However, the term *vulnerable* should be used with caution, because it also can function to stereotype or to overprotect people in some populations.[62]

Guidelines for Vulnerable Populations

In controversies over uses of vulnerable populations in biomedical research, one of three general guidelines might be applied to a research practice:

1. 1. Do not allow the practice (a policy of full prohibition).
2. 2. Allow the practice without regard to conditions (a policy of full permissibility).
3. 3. Allow the practice only under certain conditions (a policy of partial permissibility).

As an example, public opinion is deeply divided over which of these three guidelines should govern various uses of human fetuses in research—in utero and after deliberate abortions. Many prefer the first, many the second, and many the third. Divided opinions also mark debates about experimentation with animals, nontherapeutic experimentation with children, and experimentation with incompetent individuals. Few today defend either full prohibition or full permissibility of research involving these groups, but many would support a prohibition on the use of some classes of these individuals in research, including the great apes and seriously ill children. To reject the first two guidelines—as is common for some vulnerable populations—is to accept the third, which in turn requires that we establish a reasonably precise set of moral protections that fix the conditions that allow us to proceed or not to proceed with the members of a specified population.

Problems of moral coherence bedevil these issues. Near-universal agreement exists that humans who lack certain capacities should not be used in biomedical research that carries significant risk and does not offer them a prospect of direct benefit. Protections for these vulnerable populations should be at a high level because of their vulnerability. Nonhuman animals are usually not treated equivalently, though the reasons for this differential treatment are generally left unclear in public policy. Their limited cognitive and moral capacities have traditionally provided part of the substantive justification for, rather than against, their use in biomedical research when human subjects cannot ethically be used. Whether causing harm and premature death to these animals can be justified, but not justified for humans with similarly limited capacities, is an unresolved issue in biomedical ethics, and one that threatens coherence in moral theory.[63]

Practices of abortion, notably where human fetuses are capable of sentience, raise related issues of moral coherence. The long and continuing struggle over abortion primarily concerns two questions: (1) What is the

moral status of the fetus (at various developmental points)? (2) What should we do when the rights generated by this status conflict with the rights of women to control their futures? Near-universal agreement exists that an exceedingly late-term fetus is not relevantly different from a newborn. Another month earlier in development will show little in the way of morally relevant differences, and incoherence threatens any point selected on the continuum of growth as the marker of moral status. As with animal subjects, the status of human fetuses tends to be downgraded because of their lack of sentient, cognitive, and moral capacities, and this deficiency then plays a role in attempts to justify abortion. Questions about whether we can justify such downgrading and whether we can justify causing premature death to the fetus remain among the most difficult questions in biomedical ethics.

Sympathy and Impartiality

Problems of moral status and vulnerable populations raise questions about our capacity to sympathize with the predicament of others while maintaining appropriate impartiality in our judgments. In previous sections of this chapter we connected our reflections on moral status to our discussion of *moral norms* in Chapter 1. We will now connect our reflections to the account of *moral character* in Chapter 2. In particular, we focus on moral sympathy as a trait similar to compassion and usually involving empathy.

The capacity for sympathy enables us to enter into, however imperfectly, the thoughts and feelings of another individual or group. Through sympathy, we can form a concern for the other's welfare. David Hume discerningly argued that while most human beings have only a *limited* sympathy with the plight of others, they also have some level of capacity to overcome these limits through calm, reflective judgments:

> [T]he generosity of men is very limited, and … seldom extends beyond their friends and family, or, at most, beyond their native country. … [T]ho' [our] sympathy [for others] be much fainter than our concern for ourselves, and a sympathy with persons remote from us much fainter than that with persons near and contiguous; yet we neglect all these differences in our calm judgments concerning the characters of men.[64]

After we attend to ourselves, our sympathy reaches out most naturally to our intimates, such as friends and members of our family. From there sympathy can move on to a wider, but still relatively small, group of acquaintances, such as those with whom we have the most frequent contact or in whose lives we have most heavily invested. Our sympathy with those truly remote from us, such as strangers or persons in other nations, is usually diminished by comparison to sympathy with those close to us, but it can be aroused by contact with strangers and by calm judgments about their situations.

Both *dissimilarity to* and *distance from* other persons function to limit our sympathy. People in nursing homes are often both dissimilar to and distant from other persons, as are individuals with diseases such as Lesch-Nyhan, human embryos, and animals used in research. It is more difficult for many persons to view these individuals as having a significant moral status that places demands on us and holds us accountable. Even though we know that individuals in vulnerable populations suffer, our sympathy and moral responsiveness do not come easily, especially when the individuals are hidden from our view or are of another species.

Not surprisingly, many persons among the "moral saints" and some of the "moral heroes" discussed in Chapter 2 exhibit an expanded and deeper sympathy with the plight of those who suffer. Their depth of sympathy is beyond what most of us achieve or even hold as a moral ideal. By contrast, severely limited sympathy, together with severely limited generosity, helps explain social phenomena such as child abuse, animal abuse, and the neglect of enfeebled elderly persons in some nursing homes. It is regrettable that enlarged affections are not commonplace in human interactions, but this fact is predictable given what we know about human nature.

Hume proposes to address such limited sympathy for those different from us by the deliberate exercise of impartiality in our calm judgments: "It is necessary for us, in our calm judgments and discourse … to neglect all these differences, and render our sentiments more public and social."[65] He asks us to reach out and seek a more extensive sympathy. His proposals accord with our discussion in Chapter 2 of Aristotelian "moral excellence." A morally excellent person will work both to enlarge his or her sympathy for those who suffer and to reach calm

and unbiased judgments. Hume characterizes this ideal as a "common" or "general" point of view in moral judgment. This perspective, which some philosophers have called "the moral point of view," controls for the distortions and biases created by our closeness to some individuals, and opens us up to a more extensive sympathy.[66]

This perspective could help in addressing several problems encountered in this chapter, but it would be unreasonable to insist on a moral point of view that incorporates such a profoundly deep sympathy and extensive impartiality that it applies equally across cultures, populations, geography, and species. Extensive sympathy is a regulative, but arduous, ideal of conduct—as is the entire range of moral excellence examined in Chapter 2. When consistently achieved across a lifetime, it is a morally beautiful adornment of character, however rare.

CONCLUSION

In this chapter the language of "theories," "criteria," "guidelines," and "degrees" of moral status has dominated, rather than the language of "principles," "rules," "virtues," and "character" found in Chapters 1 and 2. These forms of discourse and the territories they cover should be carefully distinguished, even though they are related in various ways we have noted. For instance, the characteristics associated with moral status determine the kinds of harms and benefits an individual or group can experience. These characteristics also help to determine which moral principles apply and how they apply.

We have not argued that the common morality—as discussed in Chapters 1 and 2—gives us an adequate and workable framework of criteria of moral status, and we have left several issues about moral status undecided. There is justified uncertainty in arguments about the moral status of embryos, fetuses, brain-damaged humans, and animals used in research—and about how to analyze the idea of degrees of moral status. Reasoned disagreement is to be expected, but those who engage these issues need to be clear about the models they use and their defense, subjects rarely found in the literature of bioethics. If the model accepts degrees of moral status, that model needs to be stated with precision. If the model rejects degrees of moral status, that account, too, needs a more penetrating analysis than is usually provided. The goal of developing tiers and hierarchies of moral status is a demanding task, but its pursuit is essential in certain domains. We return to some of these problems near the end of Chapter 10, where we discuss both the common morality and the possibility of "moral change" in conceptions of moral status.

NOTES

1. 1. Cf. Mark H. Bernstein, *On Moral Considerability: An Essay on Who Morally Matters* (New York: Oxford University Press, 1998).
2. 2. This conceptual thesis is indebted to David DeGrazia, "Moral Status as a Matter of Degree," *Southern Journal of Philosophy* 46 (2008): 181–98, esp. 183. See further Tom L. Beauchamp and David DeGrazia, *Principles of Animal Research Ethics* (New York: Oxford University Press, 2019).
3. 3. For one examination of the broad range of issues involved in assessments of moral status, see the essays in *Is this Cell a Human Being? Exploring the Status of Embryos, Stem Cells and Human-Animal Hybrids*, ed. Antoine Suarez and Joachim Huarte (Germany: Springer, 2011).
4. 4. This history and its relevance for biomedical ethics are presented in Ronald A. Lindsay, "Slaves, Embryos, and Nonhuman Animals: Moral Status and the Limitations of Common Morality Theory," *Kennedy Institute of Ethics Journal* 15 (December 2005): 323–46. On the history of problems about moral status for nonhuman animals, see the four chapters by Stephen R. L. Clark, Aaron Garrett, Michael Tooley, and Sarah Chan and John Harris in *The Oxford Handbook of Animal Ethics*, ed. Tom L. Beauchamp and R. G. Frey (New York: Oxford University Press, 2011), chaps. 1–2, 11–12.
5. 5. D. J. Powner and I. M. Bernstein, "Extended Somatic Support for Pregnant Women after Brain Death," *Critical Care Medicine* 31 (2003): 1241–49; David R. Field et al., "Maternal Brain Death during Pregnancy," *JAMA: Journal of the American Medical Association* 260 (August 12, 1988): 816–22; and Xavier Bosch, "Pregnancy of Brain-Dead Mother to Continue," *Lancet* 354 (December 18–25, 1999): 2145.

6. 6. See Hilde Lindemann Nelson, "The Architect and the Bee: Some Reflections on Postmortem Pregnancy," *Bioethics* 8 (1994): 247–67; Daniel Sperling, "From the Dead to the Unborn: Is There an Ethical Duty to Save Life?" *Medicine and Law Journal* 23 (2004): 567–86; Christoph Anstotz, "Should a Brain-Dead Pregnant Woman Carry Her Child to Full Term? The Case of the 'Erlanger Baby,'" *Bioethics* 7 (1993): 340–50; and Neda Farshbaf, "Young Mother Kept Alive for 123 Days so Her Babies Could Survive," *USA Today*, July 11, 2017, available at https://www.usatoday.com/story/news/humankind/2017/07/11/young-mother-kept-alive-123-days-so-her-babies-could-survive/103615364/ (accessed April 1, 2018).
7. 7. Daniel Sperling, *Management of Post-Mortem Pregnancy: Legal and Philosophical Aspects* (Aldershot, UK: Ashgate, 2006) (addressing questions of both the moral and the legal status of the fetus); and Sarah Elliston, "Life after Death? Legal and Ethical Considerations of Maintaining Pregnancy in Brain-Dead Women," in *Intersections: Women on Law, Medicine and Technology*, ed. Kerry Petersen (Aldershot, UK: Ashgate, 1997), pp. 145–65. Our discussion does not presume that dead persons have legally protected interests and rights; we are focusing on a case in which the dead pregnant woman had an advance directive requesting that all medical technology be withheld or withdrawn under conditions that included her death.
8. 8. On this distinction, see Mary Midgley, "Duties Concerning Islands," in *Environmental Ethics*, ed. Robert Elliott (Oxford: Oxford University Press, 1995); Christopher W. Morris, "The Idea of Moral Standing," in *Oxford Handbook of Animal Ethics* (2011), pp. 261–62; and David Copp, "Animals, Fundamental Moral Standing, and Speciesism," in *Oxford Handbook of Animal Ethics* (2011), pp. 276–77.
9. 9. On why something counts "in its own right," see Allen Buchanan, "Moral Status and Human Enhancement," *Philosophy & Public Affairs* 37 (2009): 346–81, esp. 346; Frances M. Kamm, "Moral Status," in *Intricate Ethics: Rights, Responsibilities, and Permissible Harm* (New York: Oxford University Press, 2006), pp. 227–30; and L. Wayne Sumner, "A Third Way," in *The Problem of Abortion*, 3rd ed., ed. Susan Dwyer and Joel Feinberg (Belmont, CA: Wadsworth, 1997), p. 99. We thank Chris Morris for these references.
10. 10. Robert P. George and Alfonso Gómez-Lobo, "The Moral Status of the Human Embryo," *Perspectives in Biology and Medicine* 48 (2005): 201–10, quotation spanning pp. 201–5.
11. 11. Cf. the Preamble and Articles in United Nations, *Universal Declaration of Human Rights*, available at http://www.un.org/Overview/rights.html (accessed April 5, 2018).
12. 12. On September 7, 2001, V. Ourednik et al. published an article entitled "Segregation of Human Neural Stem Cells in the Developing Primate Forebrain," *Science* 293 (2001): 1820–24. This article is the first report of the implanting of human neural stem cells into the brains of a primate, creating a monkey–human chimera. The article stimulated interest in both biomedical ethics and biomedical sciences. See further National Institutes of Health (NIH), Final "National Institutes of Health Guidelines for Human Stem Cell Research" (2009). Available at https://stemcells.nih.gov/policy/2009-guidelines.htm (accessed April 5, 2018). These guidelines implement Executive Order 13505 issued on March 9, 2009, by then US President Barack Obama.
13. 13. "Chimeric" usually refers to the cellular level, whereas "transgenic" concerns the genetic level. See the argument in Mark K. Greene et al., "Moral Issues of Human–Non-Human Primate Neural Grafting," *Science* 309 (July 15, 2005): 385–86. See also the conclusions of Julian Savulescu, "Genetically Modified Animals: Should There Be Limits to Engineering the Animal Kingdom?" in *Oxford Handbook of Animal Ethics* (2011), esp. pp. 644–64; Jason Robert and Françoise Baylis, "Crossing Species Boundaries," *American Journal of Bioethics* 3 (2003): 1–13 (with commentaries); Henry T. Greely, "Defining Chimeras … and Chimeric Concerns," *American Journal of Bioethics* 3 (2003): 17–20; Robert Streiffer, "At the Edge of Humanity: Human Stem Cells, Chimeras, and Moral Status," *Kennedy Institute of Ethics Journal* 15 (2005): 347–70; and Phillip Karpowicz, Cynthia B. Cohen, and Derek van der Kooy1, "Is It Ethical to Transplant Human Stem Cells into Nonhuman Embryos?" *Nature Medicine* 10 (2004): 331–35.
14. 14. Hiromitsu Nakauchi et al., "Generation of Rat Pancreas in Mouse by Interspecific Blastocyst Injection of Pluripotent Stem Cells," *Cell* 142 (2010): 787–99. The roles of rat and mouse were reversed (i.e., swapped) in later work by this team: see T. Yamaguchi, H. Sato, M. Kato-Itoh et al., "Interspecies Organogenesis Generates Autologous Functional Islets," *Nature* 542 (2017): 191–96.
15. 15. Jun Wu, Aida Platero-Luengo, Masahiro Sakurai, et al., "Interspecies Chimerism with Mammalian Pluripotent Stem Cells," *Cell* 168 (2017): 473–86.

16. 16. National Institutes of Health (NIH), "NIH Research Involving Introduction of Human Pluripotent Cells into Non-Human Vertebrate Animal Pre-Gastrulation Embryos," Notice Number NOT-OD-15-158, Release Date September 23, 2015, available at https://grants.nih.gov/grants/guide/notice-files/NOT-OD-15-158.html (accessed March 25, 2018); and National Institutes of Health, Office of Science Policy, "Next Steps on Research Using Animal Embryos Containing Human Cells," August 4, 2016, available at http://osp.od.nih.gov/under-the-poliscope/2016/08/next-steps-research-using-animal-embryos-containing-human-cells (accessed April 1, 2018).
17. 17. See further Tom L. Beauchamp, "Moral Problems in the Quest for Human-Nonhuman Chimeras with Human Organs," *Journal of Medical Ethics*, forthcoming.
18. 18. One attractive view is that permitting the creation of animal–human hybrids for research purposes is defensible, as long as they are destroyed within a specified period of time. See Henry T. Greely, "Human/Nonhuman Chimeras: Assessing the Issues," in *Oxford Handbook of Animal Ethics* (2011), pp. 671–72, 676, 684–86. However, a federal ban on their creation was recommended by the President's Council on Bioethics, *Reproduction & Responsibility: The Regulation of New Biotechnologies* (Washington, DC: President's Council on Bioethics, 2004), available at http://bioethics.georgetown.edu/pcbe/ (accessed January 28, 2012). See also Scottish Council on Human Bioethics, *Embryonic, Fetal and Post-Natal Animal-Human Mixtures: An Ethical Discussion* (Edinburgh, UK: Scottish Council on Human Bioethics, 2010), "Animal-Human Mixtures" Publication Topic, available at http://www.schb.org.uk/ (accessed April 1, 2018).
19. 19. National Research Council, National Academy of Science, Committee on Guidelines for Human Embryonic Stem Cell Research, *Guidelines for Human Embryonic Stem Cell Research* (Washington, DC: National Academies Press, 2005), with Amendments 2007 available online at https://www.nap.edu/catalog/11871/2007-amendments-to-the-national-academies-guidelines-for-human-embryonic-stem-cell-research; and Mark Greene, "On the Origin of Species Notions and Their Ethical Limitations," in *Oxford Handbook of Animal Ethics* (2011), pp. 577–602.
20. 20. The language of "person" has a long history in theology, especially in Christian theological efforts to explicate the three individualities of the Trinity. On the potential of chimeras, see Greene et al., "Moral Issues of Human–Nonhuman Primate Neural Grafting."
21. 21. Julian Savulescu, "Should a Human-Pig Chimera Be Treated as a Person?" *Quartz*, Penned Pals, March 24, 2017, available at https://qz.com/940841/should-a-human-pig-chimera-be-treated-as-a-person/ (accessed April 5, 2017). Italics added.
22. 22. Our objections do not apply to metaphysical accounts of the nature of persons that have nothing to do with moral status. In the metaphysical literature, see Derek Parfit, "Persons, Bodies, and Human Beings," in *Contemporary Debates in Metaphysics*, ed. Theodore Sider, John Hawthorne, and Dean W. Zimmerman (Oxford: Blackwell, 2008), pp. 177–208; and Paul F. Snowdon, *Persons, Animals, Ourselves* (Oxford: Oxford University Press, 2014).
23. 23. See further Tom L. Beauchamp, "The Failure of Theories of Personhood," *Kennedy Institute of Ethics Journal* 9 (1999): 309–24; and Lisa Bartolotti, "Disputes over Moral Status: Philosophy and Science in the Future of Bioethics," *Health Care Analysis* 15 (2007): 153–58, esp. 155–57.
24. 24. At least one adherent of the first theory reaches precisely this conclusion. See Patrick Lee, "Personhood, the Moral Standing of the Unborn, and Abortion," *Linacre Quarterly* (May 1990): 80–89, esp. 87; and Lee, "Soul, Body and Personhood," *American Journal of Jurisprudence* 49 (2004): 87–125.
25. 25. For a variety of accounts see Michael Tooley, "Are Nonhuman Animals Persons?" in *Oxford Handbook of Animal Ethics* (2011), pp. 332–73; Harry G. Frankfurt, *Necessity, Volition, and Love* (Cambridge: Cambridge University Press, 1999), chaps. 9, 11; Mary Anne Warren, *Moral Status* (Oxford: Oxford University Press, 1997), chap. 1; H. Tristram Engelhardt, Jr., *The Foundations of Bioethics*, 2nd ed. (New York: Oxford University Press, 1996), chaps. 4, 6; and Lynne Rudder Baker, *Persons and Bodies* (Cambridge: Cambridge University Press, 2000), chaps. 4, 6.
26. 26. Korsgaard, "Kant's Formula of Humanity," in *Creating the Kingdom of Ends* (Cambridge: Cambridge University Press, 1996), pp. 110–11. See further her "Interacting with Animals: A Kantian Account," in *Oxford Handbook of Animal Ethics* (2011), pp. 91–118, esp. p. 103.
27. 27. See Tom Regan, *The Case for Animal Rights* (Berkeley: University of California Press, updated ed. 2004), pp. 178, 182–84.

28. 28. How this conclusion should be developed is debatable. It would be wrong to treat a late-stage Alzheimer patient in the way biomedical researchers often treat experimental animals, but it can be argued that we should treat primate research subjects with the same care taken in treating late-stage Alzheimer patients.
29. 29. See Korsgaard's assessment of what animals lack in "Interacting with Animals: A Kantian Account," p. 101.
30. 30. Colin Allen and Marc Bekoff, *Species of Mind: The Philosophy and Biology of Cognitive Ethology* (Cambridge, MA: MIT Press, 1997); and Colin Allen, "Assessing Animal Cognition: Ethological and Philosophical Perspectives," *Journal of Animal Science* 76 (1998): 42–47.
31. 31. See Donald R. Griffin, *Animal Minds: Beyond Cognition to Consciousness*, 2nd ed. (Chicago: University of Chicago Press, 2001); Rosemary Rodd, *Ethics, Biology, and Animals* (Oxford: Clarendon, 1990), esp. chaps. 3–4, 10; and Tom L. Beauchamp and Victoria Wobber, "Autonomy in Chimpanzees," *Theoretical Medicine and Bioethics* 35 (April 2014): 117–32.
32. 32. Cf. Gordon G. Gallup, "Self-Recognition in Primates," *American Psychologist* 32 (1977): 329–38; and David DeGrazia, *Taking Animals Seriously: Mental Life and Moral Status* (New York: Cambridge University Press, 1996), esp. p. 302.
33. 33. A full account of these criteria would require explication in terms of some of the cognitive conditions discussed previously. For example, the capacity to make moral judgments requires a certain level of the capacity for understanding.
34. 34. Kant, *Grounding for the Metaphysics of Morals*, trans. James W. Ellington, in Kant, *Ethical Philosophy* (Indianapolis, IN: Hackett, 1983), pp. 38–41, 43–44 (Preussische Akademie, pp. 432, 435, 436, 439–40).
35. 35. Examples of such theories—focused on the claim that there is sufficient evidence to count some nonhuman animals as moral agents, possibly persons, and therefore as members of the moral community—are Marc Bekoff and Jessica Pierce, *Wild Justice: The Moral Lives of Animals* (Chicago: University of Chicago Press, 2009); Steven M. Wise, *Rattling the Cage: Toward Legal Rights for Animals* (Boston: Da Capo Press of Perseus Books, 2014, updated ed.); Michael Bradie, "The Moral Life of Animals," in *Oxford Handbook of Animal Ethics* (2011), pp. 547–73, esp. pp. 555–70; and Tom Regan, *The Case for Animal Rights*, esp. pp. 151–56.
36. 36. See Colin Allen and Michael Trestman, "Animal Consciousness," *Stanford Encyclopedia of Philosophy*, substantive revision of October 24, 2016, especially sections 6–7, available at https://plato.stanford.edu/entries/consciousness-animal/ (accessed June 12, 2018); and David Edelman, Bernard Baars, and Anil Seth, "Identifying Hallmarks of Consciousness in Non-Mammalian Species," *Consciousness and Cognition* 14 (2005): 169–87.
37. 37. The terms *pain* and *suffering* are frequently used interchangeably, but they should be distinguished on grounds that suffering may require more cognitive ability than the mere experience of pain. Suffering may occur from aversive or harmful states such as misery that are not attended by pain. For a close analysis of suffering and related notions, see David DeGrazia, "What Is Suffering and What Kinds of Beings Can Suffer?" in *Suffering and Bioethics*, ed. Ronald Green and Nathan Palpant (New York: Oxford University Press, 2014): 134–53. See also Robert Elwood, "Pain and Suffering in Invertebrates?" *ILAR Journal* 52 (2011): 175–84; Tom L. Beauchamp and David B. Morton, "The Upper Limits of Pain and Suffering in Animal Research: A Moral Assessment of The European Union's Legislative Framework," *Cambridge Quarterly of Healthcare Ethics* 24 (October 2015): 431–47; and David DeGrazia and Tom L. Beauchamp "Moving Beyond the Three Rs," *ILAR Journal* 61 (Fall 2019).
38. 38. Some defenders also seem to claim that this capacity is both *necessary and sufficient* for moral status—a more difficult claim to support. See two opposed theories on this issue in L. Wayne Sumner, *Abortion and Moral Theory* (Princeton, NJ: Princeton University Press, 1981); and Bonnie Steinbock, *Life before Birth: The Moral and Legal Status of Embryos and Fetuses*, 2nd ed. (New York: Oxford University Press, 2011).
39. 39. Baruch Brody, *Abortion and the Sanctity of Life* (Cambridge, MA: MIT Press, 1975). Brain birth is said to be analogous to brain death at critical transition points.
40. 40. This point is made in Stephen Griffith, "Fetal Death, Fetal Pain, and the Moral Standing of a Fetus," *Public Affairs Quarterly* 9 (1995): 117.

41. 41. Bentham, *An Introduction to the Principles of Morals and Legislation*, ed. J. H. Burns and H. L. A. Hart; with a new introduction by F. Rosen; and an interpretive essay by Hart (Oxford: Clarendon Press, 1996), p. 283.
42. 42. See, for example, Peter Singer, *Animal Liberation*, 2nd ed. (London: Pimlico, 1995), p. 8; and Sumner, *Abortion and Moral Theory.*
43. 43. See R. G. Frey, "Moral Standing, the Value of Lives, and Speciesism," *Between the Species* 4 (Summer 1988): 191–201; "Animals," in *The Oxford Handbook of Practical Ethics* (New York: Oxford University Press, 2003), esp. pp. 163, 178; and his "Autonomy and the Value of Animal Life," *Monist* 70 (January 1987): 50–63. A somewhat similar, but differently grounded, theory appears in Martha Nussbaum, *Frontiers of Justice: Disability, Nationality, Species Membership* (Cambridge, MA: Harvard University Press, 2006), especially p. 361.
44. 44. For relevant theoretical literature, see Ronald M. Green, "Determining Moral Status," *American Journal of Bioethics* 2 (Winter 2002): 20–30; and Diane Jeske, "Special Obligations," *Stanford Encyclopedia of Philosophy* (Spring 2014 Edition), ed. Edward N. Zalta, available at https://plato.stanford.edu/archives/spr2014/entries/special-obligations/ (accessed March 28, 2018). For a compelling account of how bonding can occur with animal research subjects and its moral importance, see John P. Gluck, *Voracious Science and Vulnerable Animals: A Primate Scientist's Ethical Journey* (Chicago: University of Chicago Press, 2016); and see also Lily-Marlene Russow, "Ethical Implications of the Human-Animal Bond in the Laboratory," *ILAR Journal* 43 (2002): 33–37.
45. 45. Carson Strong and Garland Anderson, "The Moral Status of the Near-Term Fetus," *Journal of Medical Ethics* 15 (1989): 25–26.
46. 46. See the related conclusion in Nancy Jecker, "The Moral Status of Patients Who Are Not Strict Persons," *Journal of Clinical Ethics* 1 (1990): 35–38.
47. 47. For a broader set of patients than this list suggests—especially countless terminally ill patients—see Felicia Cohn and Joanne Lynn, "Vulnerable People: Practical Rejoinders to Claims in Favor of Assisted Suicide," in *The Case against Assisted Suicide: For the Right to End-of-Life Care*, ed. Kathleen Foley and Herbert Hendin (Baltimore: Johns Hopkins University Press, 2002), pp. 238–60.
48. 48. An influential general strategy of melding diverse theories is proposed in Warren, *Moral Status*, though her set of melded theories differs from ours. A similar strategy, with a different set of melded theories, appears in Lawrence J. Nelson and Michael J. Meyer, "Confronting Deep Moral Disagreements: The President's Council on Bioethics, Moral Status, and Human Embryos," *American Journal of Bioethics* 5 (2005): 33–42 (with a response to critics, pp. W14–16).
49. 49. The problem of equal and unequal consideration of interests, and different degrees of consideration, is discussed in DeGrazia, "Moral Status as a Matter of Degree," esp. pp. 188, 191.
50. 50. [Mary Warnock], *Report of the Committee of Inquiry into Human Fertilisation and Embryology: Presented to Parliament* (London: HMSO, July 1984). [The Warnock Committee Report.]
51. 51. Chief Medical Officer's Expert Group, *Stem Cell Research: Medical Progress with Responsibility* (London: Department of Health, 2000).
52. 52. Chief Medical Officer's Expert Group, *Stem Cell Research*, sects. 4.6, 4.12, pp. 38–39.
53. 53. See David DeGrazia, "Great Apes, Dolphins, and the Concept of Personhood," *Southern Journal of Philosophy* 35 (1997): 301–20; and Beauchamp, "The Failure of Theories of Personhood."
54. 54. For an all-or-nothing account that rejects degrees of moral status, see Elizabeth Harman, "The Potentiality Problem," *Philosophical Studies* 114 (2003): 173–98.
55. 55. Carson Strong, "The Moral Status of Preembryos, Embryos, Fetuses, and Infants," *Journal of Medicine and Philosophy* 22 (1997): 457–78.
56. 56. Cf. the similar conclusion, with an argued defense, in Mary Anne Warren, "Moral Status," in *A Companion to Applied Ethics*, ed. R. G. Frey and Christopher Wellman (Oxford: Blackwell, 2003), p. 163. See further Elizabeth Harman, "Creation Ethics: The Moral Status of Early Fetuses and the Ethics of Abortion," *Philosophy & Public Affairs* 28 (1999): 310–324.
57. 57. For related, yet different, objections to this account, see Rebecca L. Walker, "Beyond Primates: Research Protections and Animal Moral Value," *Hastings Center Report* 46 (2016): 28–30.
58. 58. See Mary Midgley, *Animals and Why They Matter* (Athens: University of Georgia Press, 1983), pp. 28–30, 100; Rosalind Hursthouse, "Virtue Ethics and the Treatment of Animals," in *Oxford Handbook of*

Animal Ethics (2011), chap. 4; and Hursthouse, *Ethics, Humans and Other Animals* (London: Routledge, 2000), pp. 127–32.

59. 59. Classic cases in the United States are the Tuskegee syphilis experiment, the use of children with intellectual disabilities at the Willowbrook State School, and the injection of cancer cells into debilitated patients at the Jewish Chronic Disease Hospital in Brooklyn. For the first, see James H. Jones, *Bad Blood: The Tuskegee Syphilis Experiment*, rev. ed. (New York: Free Press, 1993), and Susan Reverby, ed., *Tuskegee's Truths: Rethinking the Tuskegee Syphilis Study* (Chapel Hill: University of North Carolina Press, 2000). For the others, see Jay Katz et al., eds., *Experimentation with Human Beings: The Authority of the Investigator, Subject, Professions, and State in the Human Experimentation Process* (New York: Russell Sage Foundation, 1972); and National Commission for the Protection of Human Subjects of Biomedical and Behavioral Research, *Research Involving Those Institutionalized as Mentally Infirm* (Washington: Department of Health, Education, and Welfare [DHEW], 1978).
60. 60. Parallel debates in environmental ethics focus on the moral status of dimensions of nature beyond human and nonhuman animals; for example, whether individual trees, plants, species, and ecosystems have moral status. See Paul Taylor, *Respect for Nature: A Theory of Environmental Ethics* (Princeton, NJ: Princeton University Press, 2011); Gary Varner, "Environmental Ethics, Hunting, and the Place of Animals," *Oxford Handbook of Animal Ethics* (2011), pp. 855–76; Andrew Brennan and Y. S. Lo, *Understanding Environmental Philosophy* (New York: Routledge, 2014); Lawrence E. Johnson, *A Morally Deep World: An Essay on Moral Significance and Environmental Ethics* (Cambridge: Cambridge University Press, 1993); Agnieszka Jaworska and Julie Tannenbaum, "The Grounds of Moral Status," *Stanford Encyclopedia of Philosophy* (revision of January 10, 2018), available at https://plato.stanford.edu/entries/grounds-moral-status/ (accessed March 19, 2018); and Alasdair Cochrane, "Environmental Ethics," section 1 ("Moral Standing"), *Internet Encyclopedia of Philosophy*, available at https://www.iep.utm.edu/envi-eth/ (accessed March 19, 2018).
61. 61. National Commission for the Protection of Human Subjects of Biomedical and Behavioral Research, *The Belmont Report: Ethical Principles and Guidelines for the Protection of Human Subjects of Research* (Washington, DC: DHEW Publication OS 78–0012, 1978); Code of Federal Regulations, Title 45 (Public Welfare), Part 46 (Protection of Human Subjects), http://www.hhs.gov/ohrp/humansubjects/guidance/45cfr46.html (accessed July 15, 2011).
62. 62. On analysis of vulnerability, see Kenneth Kipnis, "Vulnerability in Research Subjects: A Bioethical Taxonomy," in National Bioethics Advisory Commission (NBAC), *Ethical and Policy Issues in Research Involving Human Participants*, vol. 2 (Bethesda, MD: NBAC, 2001), pp. G-1–13.
63. 63. See Rebecca L. Walker, "Human and Animal Subjects of Research: The Moral Significance of Respect versus Welfare," *Theoretical Medicine and Bioe*thics 27 (2006): 305–31. A major document that illustrates the problem is an Institute of Medicine (now National Academy of Medicine) report: Committee on the Use of Chimpanzees in Biomedical and Behavioral Research, *Chimpanzees in Biomedical and Behavioral Research: Assessing the Necessity* (Washington, DC: National Academies Press, 2011), available at https://www.nap.edu/catalog/13257/chimpanzees-in-biomedical-and-behavioral-research-assessing-the-necessity (retrieved August 16, 2017). See also National Institutes of Health, Office of the Director, "Statement by NIH Director Dr. Francis Collins on the Institute of Medicine Report Addressing the Scientific Need for the Use of Chimpanzees in Research," Thursday, December 15, 2011, available at http://www.nih.gov/news/health/dec2011/od-15.htm (accessed December 15, 2011); and the follow-up report, Council of Councils, National Institutes of Health. *Council of Councils Working Group on the Use of Chimpanzees in NIH-Supported Research: Report*, 2013, available at https://dpcpsi.nih.gov/council/pdf/FNL_Report_WG_Chimpanzees.pdf (accessed August 16, 2017); National Institutes of Health, Announcement of Agency Decision: Recommendations on the Use of Chimpanzees in NIH-Supported Research, available at dpcpsi.nih.gov/council/pdf/NIHresponse_to_Council_of_Councils_recommendations_62513.pdf (accessed July 28, 2013).
64. 64. Hume, *A Treatise of Human Nature*, ed. David Fate Norton and Mary J. Norton (Oxford: Oxford University Press, 2006), 3.3.3.2.
65. 65. Hume, *An Enquiry Concerning the Principles of Morals*, ed. Tom L. Beauchamp (Oxford: Oxford University Press, 1998), 5.42.

66. 66. We here concentrate on the role impartiality plays in expanding sympathy, but impartiality also can help correct misdirected and exaggerated sympathy that borders on sentimentality. For a critique of a kind of sentimentality that stands opposed to potentially effective measures to obtain transplantable organs from brain-dead individuals, see Joel Feinberg, “The Mistreatment of Dead Bodies,” *Hastings Center Report* 15 (February 1985): 31–37.

4

Respect for Autonomy

The principle of respect for the autonomous choices of persons runs as deep in morality as any principle, but determining its nature, scope, and strength requires careful analysis. We explore the concept of autonomy and the principle of respect for autonomy in this chapter primarily to examine patients', subjects', and surrogates' decision making in health care and research.[1]

We begin our analysis of a framework of four principles of biomedical ethics with this principle of respect, but the order of our chapters does not imply that this principle has moral priority over, or a more foundational status than, other principles. Not only do we hold that the principle of respect for autonomy lacks priority over the other principles, but we maintain that it is not excessively individualistic to the neglect of the social nature of individuals, not excessively focused on reason to the neglect of the emotions, and not unduly legalistic by highlighting legal rights while downplaying social practices.

The Concept of Autonomy and the Principle of Respect for Autonomy

The word *autonomy*, derived from the Greek *autos* ("self") and *nomos* ("rule," "governance," or "law"), originally referred to the self-rule or self-governance of independent city-states. Autonomy has since been extended to individuals. The autonomous individual acts freely in accordance with a self-chosen plan, analogous to the way an autonomous government manages its territories and sets its policies. In contrast, a person of diminished autonomy is substantially controlled by others or incapable of deliberating or acting on the basis of his or her desires and plans. For example, cognitively impaired individuals and prisoners often have diminished autonomy. Mental incapacitation limits the autonomy of a person with a severe mental handicap, and incarceration constrains a prisoner's autonomy.

Two general conditions are essential for autonomy: *liberty* (independence from controlling influences) and *agency* (capacity for intentional action). However, disagreement exists over the precise meaning of these two conditions and over whether additional conditions are required for autonomy.[2] As our first order of business, we use these basic conditions to construct a theory of autonomy that we believe suitable for biomedical ethics.

Theories of Autonomy

Some theories of autonomy feature the abilities, skills, or traits of the *autonomous person*, which include capacities of self-governance such as understanding, reasoning, deliberating, managing, and independent choosing.[3] Our focus in this chapter on decision making leads us to concentrate on *autonomous choice* rather than general capacities for self-governance and self-management. Even autonomous persons who have self-governing capacities, and generally manage their health well, sometimes fail to govern themselves in particular choices because of temporary constraints caused by illness, depression, ignorance, coercion, or other conditions that limit their judgment or their options.

An autonomous person who signs a consent form for a procedure without reading or understanding the form has the capacity to act autonomously but fails to so act in this circumstance. Depending on the context, we might be able to correctly describe the act as that of placing trust in one's physician and therefore as an act that autonomously authorizes the physician to proceed. However, even if this claim is accurate, the act is not an autonomous authorization *of the procedure* because this person lacks material information about the procedure. Similarly, some persons who are generally incapable of autonomous decision making can at times make autonomous choices. For example, some patients in mental institutions who cannot care for themselves and have

been declared legally incompetent may still be competent to make some autonomous choices, such as stating preferences for meals, refusing some medications, and making phone calls to acquaintances.

Split-level theories of autonomy. Some philosophers have presented an influential theory of autonomy that requires having the capacity to reflectively control and identify with or oppose one's basic (first-order) desires or preferences through higher level (second-order) desires or preferences.[4] Gerald Dworkin, for instance, offers a "content-free" definition of autonomy as a "second-order capacity of persons to reflect critically upon their first-order preferences, desires, wishes, and so forth and the capacity to accept or attempt to change these in the light of higher-order preferences and values."[5] An example is an alcoholic who has a desire to drink but also has a higher-order desire to stop drinking. A second example is an exceptionally dedicated physician who has a first-order desire to work extraordinarily long hours in the hospital while also having a higher-order commitment to spend all of her evening hours with her family. Whenever she wants to work late in the evening and does so, she wants what she does not autonomously want, and therefore acts nonautonomously. Action from a first-order desire that is not endorsed by a second-order volition is not autonomous and represents "animal" behavior. Accordingly, in this theory an autonomous person is one who has the capacity to reflectively accept, identify with, or repudiate a lower-order desire independent of others' manipulations of that desire. This higher-order capacity to accept or repudiate first-order preferences *constitutes* autonomy, and no person is autonomous without this capacity.

This theory is problematic because nothing prevents a reflective acceptance, preference, or volition at the second level from being caused by a strong first-order desire. That is, the individual's second-level acceptance of a first-order desire may be the causal result of an already formed structure of first-order preferences. Potent first-order desires from a condition such as alcohol or opioid addiction are antithetical to autonomy and can cause second-order desires. If second-order desires (decisions, volitions, etc.) are generated by first-order desires, then the process of identifying with one desire rather than another does not distinguish autonomy from nonautonomy.

This theory needs more than a convincing account of second-order preferences and acceptable influences. It needs a way for ordinary persons to qualify as deserving respect for their autonomous choices even when they have not reflected on their preferences at a higher level. The theory also risks running afoul of the criterion of coherence with the principle of respect for autonomy discussed throughout this chapter. If reflective identification with one's second-order desires or volitions is a necessary condition of autonomous action, then many ordinary actions that are almost universally considered autonomous, such as cheating on one's spouse (when one truly wishes not to be such a person) or selecting tasty snack foods when grocery shopping (when one has never reflected on one's desires for snack foods), would be *non*autonomous in this theory. A theory that requires reflective identification and stable volitional patterns unduly narrows the scope of actions protected by the principle of respect for autonomy.

Agnieszka Jaworska insightfully argues that choosing contrary to one's professed, accepted, and stable values need not constitute an abandonment of autonomy. For example, a patient might request a highly invasive treatment at the end of life against his previous convictions about his best interests because he has come to a conclusion that surprises him: He cares more about living a few extra days than he had thought he would. Despite his long-standing and firm view that he would reject such invasive treatments, he now accepts them. Jaworska's example is common in medical contexts.[6]

Few decision makers and few choices would be autonomous if held to the standards of higher-order reflection demanded by this split-level theory. It presents an aspirational ideal of autonomy rather than a theory of autonomy suitable for decision making in health care and research. A theory should not be inconsistent with pretheoretical assumptions implicit in the principle of respect for autonomy, and no theory of autonomy is acceptable if it presents an ideal beyond the reach of competent choosers.

Our three-condition theory. Instead of an ideal theory of autonomy, our analysis focuses on nonideal conditions. We analyze autonomous action in terms of normal choosers who act (1) intentionally, (2) with understanding, and (3) without controlling influences that determine their action. This uncomplicated account is designed to be

coherent with the premise that the everyday choices of generally competent persons are autonomous and to be sufficient as an account of autonomy for biomedical ethics.

1. 1. *Intentionality.* Intentional actions require plans in the form of representations of the series of events proposed for the execution of an action. For an act to be intentional it must correspond to the actor's conception of the act in question, although a planned outcome might not materialize as projected.[7] Nothing about intentional acts rules out actions that an agent wishes he or she did not have to perform. Our motivation often involves *conflicting* wants and desires, but this fact does not render an action less than intentional or autonomous. Foreseen but undesired outcomes can be part of a coherent plan of intentional action.
2. 2. *Understanding.* Understanding is the second condition of autonomous action. An action is not autonomous if the actor does not adequately understand it. Conditions that limit understanding include illness, irrationality, and immaturity. Deficiencies in a communication process also can hamper understanding. An autonomous action needs only a substantial degree of understanding, not a full understanding. To restrict adequate decision making by patients and research subjects to the ideal of fully or completely autonomous decision making strips their acts of a meaningful place in the practical world, where people's actions are rarely, if ever, fully autonomous.
3. 3. *Noncontrol.* The third of the three conditions of autonomous action is that a person be free of controls exerted either by external sources or by internal states that rob the person of self-directedness. Influence and resistance to influence are basic concepts in this analysis. Not all influences exerted on another person are controlling. Our analysis of noncontrol and voluntariness later in this chapter focuses on coercion and manipulation as key categories of influence. We concentrate on *external* controlling influences—usually influences of one person on another—but no less important to autonomy are *internal* influences on the person, such as those caused by mental illness.

The first of the three conditions of autonomy—intentionality—is not a matter of degree: Acts are either intentional or nonintentional. However, acts can satisfy the conditions of both understanding and absence of controlling influence to a greater or lesser extent. For example, understanding can be more or less complete; threats can be more or less severe; and mental illness can be more or less controlling. Children provide a good example of the continuum from being in control to not being in control. In the early months of life children are heavily controlled and display only limited ability to exercise control: They exhibit different degrees of resistance to influence as they mature, and their capacity to take control and perform intentional actions, as well as to understand, gradually increases.

Acts therefore can be autonomous by degrees, as a function of satisfying these two conditions of understanding and voluntariness to different degrees. A continuum of both understanding and noncontrol runs from full understanding and being entirely in control to total absence of relevant understanding and being fully controlled. Cutoff points on these continua are required for the classification of an action as either autonomous or nonautonomous. The lines between adequate and inadequate degrees of understanding and degrees of control must be determined in light of specific objectives of decision making in a particular context such as deciding about surgery, choosing a university to attend, and hiring a new employee.

Although the line between what is substantial and what is insubstantial may appear arbitrary, thresholds marking substantially autonomous decisions can be appropriately set in light of specific objectives of decision making. Patients and research subjects can achieve substantial autonomy in their decisions, just as substantially autonomous choice occurs in other areas of life, such as selecting a diet. We need to formulate specific criteria for substantial autonomy in particular contexts.

Autonomy, Authority, Community, and Relationships

Some theorists argue that autonomous action is incompatible with the authority of governments, religious organizations, and other communities that prescribe behavior. They maintain that autonomous persons must act on their own reasons and cannot submit to an authority or choose to be ruled by others without relinquishing their autonomy.[8] However, no fundamental inconsistency exists between autonomy and authority if individuals

exercise their autonomy to choose to accept an institution, tradition, or community that they view as a legitimate source of influence and direction.

Choosing to strictly follow the recommendations of a medical authority is a prime example. Other examples are a Jehovah's Witness who accepts the authority of that tradition and refuses a recommended blood transfusion or a Roman Catholic who chooses against an abortion in deference to the authority of the church. That persons share moral norms with authoritative institutions does not prevent these norms from being autonomously accepted, even if the norms derive from traditions or from institutional authority. If a Jehovah's Witness who insists on adhering to the doctrines of his faith in refusing a blood transfusion is deemed nonautonomous on the basis of his religious convictions, many of our choices based on our confidence in institutional authority will be likewise deemed unworthy of respect. A theory of autonomy that makes such a demand is morally unacceptable.

We encounter many limitations of autonomous choice in medical contexts because of the patient's dependent condition and the medical professional's authoritative position. On some occasions authority and autonomy are not compatible, but this is not because the two *concepts* are incompatible. Conflict may arise because authority has not been properly presented or accepted, as in certain forms of medical paternalism or when an undue influence has been exerted.

Some critics of autonomy's prominent role in biomedical ethics question what they deem to be a model of an independent, rational will inattentive to emotions, communal life, social context, interdependence, reciprocity, and the development of persons over time. They see such an account of autonomy as too narrowly focused on the self as independent, atomistic, and rationally controlling. Some of these critics have sought to affirm autonomy while interpreting it through relationships.[9] This account of "relational autonomy" is motivated by the conviction that persons' identities and choices are generally shaped, for better or worse, through social interactions and intersecting social determinants such as race, class, gender, ethnicity, and authority structures.[10]

We will address the challenges of relational autonomy through the ethical principles analyzed in Chapters 5 through 7. In our view, a relational conception of autonomy can be defensible if it does not neglect or obscure the three conditions of autonomy we identified previously and will further analyze later in this chapter.

The Principle of Respect for Autonomy

To respect autonomous agents is to acknowledge their right to hold views, to make choices, and to take actions based on their values and beliefs. Respect is shown through respectful *action*, not merely by a respectful *attitude*. The principle of respect for autonomy requires more than noninterference in others' personal affairs. In some contexts it includes building up or maintaining others' capacities for autonomous choice while helping to allay fears and other conditions that destroy or disrupt autonomous action. Respect involves acknowledging the value and decision-making rights of autonomous persons and enabling them to act autonomously, whereas disrespect for autonomy involves attitudes and actions that ignore, insult, demean, or are inattentive to others' rights of autonomous action.

The principle of respect for autonomy asserts a broad obligation that is free of exceptive clauses such as "We must respect individuals' views and rights *except when* their thoughts and actions seriously harm other persons." Exceptive conditions should appear in specifications of the principle, not in the principle itself. However, the principle should be analyzed as containing both a negative obligation and a positive obligation. As a *negative* obligation, the principle requires that autonomous actions not be subjected to controlling constraints by others. As a *positive* obligation, the principle requires both respectful disclosures of information and other actions that foster autonomous decision making. Respect for autonomy obligates professionals in health care and research involving human subjects to disclose information, to probe for and ensure understanding and voluntariness, and to foster adequate decision making. As some contemporary Kantians have appropriately pointed out, the moral demand that we treat others as ends requires that we assist them in achieving their ends and foster their capacities as agents, not merely that we avoid treating them solely as means to our ends.[11]

These negative and positive sides of respect for autonomy support more specific moral rules, some of which may also be justified, in part, by other moral principles discussed in this book. Examples of these rules include the following:

1. 1. Tell the truth.
2. 2. Respect the privacy of others.
3. 3. Protect confidential information.
4. 4. Obtain consent for interventions with patients.
5. 5. When asked, help others make important decisions.

The principle of respect for autonomy and each of these rules has only prima facie standing, and competing moral considerations sometimes override them. Examples include the following: If our autonomous choices endanger the public health, potentially harm innocent others, or require a scarce resource for which no funds are available, others can justifiably restrict our exercises of autonomy. The principle of respect for autonomy often does not determine what, on balance, a person ought to be free to know or do or what counts as a valid justification for constraining autonomy. For example, a patient with an inoperable, incurable carcinoma once asked, "I don't have cancer, do I?" The physician lied, saying, "You're as good as you were ten years ago." This lie infringed the principle of respect for autonomy by denying the patient information he may have needed to determine his future courses of action. Although the matter is controversial, such a lie might be justified by a principle of beneficence if *major* benefits will flow to the patient. (For the justification of certain acts of withholding the truth from patients, see our discussions of paternalism in Chapter 6 and veracity in Chapter 8.)

Obligations to respect autonomy do not extend to persons who cannot act in a sufficiently autonomous manner and to those who cannot be rendered autonomous because they are immature, incapacitated, ignorant, coerced, exploited, or the like. Infants, irrationally suicidal individuals, and drug-dependent patients are examples. This standpoint does not presume that these individuals are not owed moral respect, often referred to as respect for persons.[12] In several of our chapters we show that these patients have a significant moral status (see Chapter 3) that obligates us to protect them from harm-causing conditions and to supply medical benefits to them (see Chapters 5–7).

The Alleged Triumph and Failure of Respect for Autonomy

Some writers lament the "triumph of autonomy" in American bioethics. They assert that autonomy's proponents sometimes disrespect patients by forcing them to make choices, even though many patients do not want to receive information about their condition or to make decisions. Carl Schneider, for example, claims that stout proponents of autonomy, whom he labels "autonomists," concern themselves less with what patients *do want* than with what they *should want*. He concludes that "while patients largely wish to be informed about their medical circumstances, a substantial number of them [especially the elderly and the very sick] do not want to make their own medical decisions, or perhaps even to participate in those decisions in any very significant way."[13]

A health professional's duty of respect for autonomy correlates with the *right* of a patient or subject to choose, but the patient or subject does not have a correlative *duty* to choose. Several empirical studies of the sort cited by Schneider seem to misunderstand, as he does, how autonomous choice functions in a viable theory and how it should function in clinical medicine. In one study, UCLA researchers examined the differences in the attitudes of elderly subjects, sixty-five years old or older, from different ethnic backgrounds toward (1) disclosure of the diagnosis and prognosis of a terminal illness, and (2) decision making at the end of life. The researchers summarize their main findings, based on 800 subjects (200 from each ethnic group):

> Korean Americans (47%) and Mexican Americans (65%) were significantly less likely than European Americans (87%) and African Americans (88%) to believe that a patient should be told the diagnosis of metastatic cancer. Korean Americans (35%) and Mexican Americans (48%) were less likely than African Americans (63%) and European Americans (69%) to believe that a patient should be told of a terminal prognosis and less likely to believe that the patient should make

decisions about the use of life-supporting technology (28% and 41% vs. 60% and 65%). Korean Americans and Mexican Americans tended to believe that the family should make decisions about the use of life support.

Investigators in this study stress that "belief in the *ideal* of patient autonomy is far from universal" (italics added), and they contrast this ideal with a "family-centered model" focused on an individual's web of relationships and "the harmonious functioning of the family."[14] Nevertheless, the investigators conclude that "physicians should ask their patients if they wish to receive information and make decisions or if they prefer that their families handle such matters." Far from abandoning or supplanting the moral demand that we respect individual autonomy, their recommendation accepts the normative position that the choice is rightly the patient's or a designated surrogate's. Even if the patient delegates the right to someone else, his or her choice to delegate can be autonomous.

In a second study, this time of Navajo values and the disclosure of risk and medical prognoses, two researchers sought to determine how health care providers "should approach the discussion of negative information with Navajo patients" to provide "more culturally appropriate medical care." Frequent conflicts emerge, these researchers report, between autonomy and the traditional Navajo conception that "thought and language have the power to shape reality and to control events." In the traditional conception, telling a Navajo patient recently diagnosed with a disease the potential complications of that disease could actually produce those complications, because "language does not merely describe reality, language shapes reality." Traditional Navajo patients may process negative information as dangerous to them. They expect instead a "positive ritual language" that promotes or restores health.

One middle-aged Navajo nurse reported that a surgeon explained the risks of bypass surgery to her father in such a way that he refused to undergo the procedure: "The surgeon told him that he may not wake up, that this is the risk of every surgery. For the surgeon it was very routine, but the way that my Dad received it, it was almost like a death sentence, and he never consented to the surgery." The researchers therefore found ethically troublesome policies that attempt to "expose all hospitalized Navajo patients to the idea, if not the practice, of advance care planning."[15]

These two studies enrich our understanding of diverse cultural beliefs and values. However, these studies sometimes misrepresent what the principle of respect for autonomy and related laws and policies require. They view their results as opposing rather than, as we interpret them, enriching the principle of respect for autonomy. A fundamental obligation exists to ensure that patients have the right to choose as well as the right to accept or decline information. Forced information and forced choice are usually inconsistent with this obligation.

A tension exists between the two studies just discussed. One study recommends inquiring in advance to ascertain patients' preferences about information and decision making, whereas the other suggests, tenuously, that even informing certain patients of a right to decide may cause harm. The practical question is whether it is possible to inform patients of their rights to know and to decide without compromising their systems of belief and values or otherwise disrespecting them by forcing them to learn or choose when a better form of communication could avoid this outcome. Health professionals should almost always inquire about their patients' wishes to receive information and to make decisions and should not assume that because a patient belongs to a particular community or culture, he or she affirms that community's customary worldview and values. The main requirement is to respect a particular patient's or subject's autonomous choices, whatever they may be. Respect for autonomy is no mere ideal in health care; it is a professional obligation.

Complexities in Respecting Autonomy

Varieties of autonomous consent. Consent often grants permission for others to act in ways that are unjustifiable without consent—for instance, engaging in sexual relations or performing surgery. However, when examining autonomy and consent in this chapter, we do not presume that consent is either necessary or sufficient for certain interventions to be justified. It is not always necessary in emergencies, in public health interventions, in research involving anonymized data, and so forth; and it is not always sufficient because other ethical principles too must

be satisfied. For example, research involving human subjects must pass a benefit-risk test and a fairness test in the recruitment of participants.[16]

The basic paradigm of the exercise of autonomy in health care and in research is *express* or *explicit* consent (or refusal), usually informed consent (or refusal).[17] However, the informed consent paradigm captures only one form of valid consent. Consent may also be implied, tacit, or presumed; and it may be general or specific.

Implicit (or *implied*) consent is inferable from actions. Consent to a medical procedure may be implicit in a specific consent to another procedure, and providing general consent to treatment in a teaching hospital may imply consent to various roles for physicians, nurses, and others in training. Another form is *tacit* consent, which occurs silently or passively through omissions. For example, if the staff of a long-term care facility asks residents whether they object to having the time of dinner changed by one hour, a uniform lack of objection constitutes consent.

Presumed consent is subject to a variety of interpretations. It is a form of implicit consent if consent is presumed on the basis of what is known about a particular person's choices. In certain contexts, presumed consent is tacit consent that gives good grounds for accepting the consent as valid. By contrast, presuming consent on the basis of either a theory of human goods that are desirable or what a rational person would accept is morally perilous. Consent should refer to an individual's actual choices or known preferences, not to presumptions about the choices the individual would or should make.

Different conceptions of consent have appeared in debates about teaching medical students how to perform intimate examinations, especially pelvic and rectal examinations.[18] Medical students have often learned and practiced on anesthetized patients, some of whom have not given an explicit informed consent. For instance, some teaching hospitals have allowed one or two medical students to participate in the examination of women who are under anesthesia in preparation for surgery. Anesthetized patients have been considered ideal for teaching medical students how to perform a pelvic examination because these patients are relaxed and would not feel any mistakes. When questioned about this practice, some directors of obstetrics and gynecology programs appealed to the patient's general consent upon entering a teaching hospital. This consent typically authorizes medical students and residents to participate in patients' care for teaching and learning purposes. However, the procedures that involve participation by medical students or other medical trainees are often not explicitly stated.

There are good ethical reasons to find general consent insufficient and, instead, to require specific informed consent for such intimate examinations performed for educational or training purposes. Health professionals usually—and rightly—seek specific informed consent when a procedure is invasive, as in surgery, or when it is risky. Although pelvic examinations are not invasive or risky by comparison to surgery, patients may object to these intrusions into their bodies, especially for purposes of education and training. When asked, many women consent to the participation of medical students in such examinations, but other women view the practice as a violation of their dignity and privacy.[19] One commentator appropriately maintains that "the patient must be treated as the student's teacher, not as a training tool."[20]

Using anesthetized women who have given only a general consent may be efficient in clinical training, but, in view of the importance of respect for autonomy, it is ethically required, instead, to use only anesthetized patients who have given specific informed consent or healthy volunteers willing to serve as standardized patients. Both alternatives respect personal autonomy, avoid an inappropriate form of medical education, and are workable.[21]

The practice of conducting pelvic exams on anesthetized patients without their specific informed consent also may have a negative impact on clinicians' attitudes toward the importance of informed consent and, by implication, toward respect for autonomy. According to a study of medical students in the Philadelphia area, this practice desensitized physicians to the need for patients to give their consent before these and presumably other procedures. For students who had finished an obstetrics/gynecology clerkship, which involved this practice, consent was significantly less important (51%) than for students who had not completed a clerkship (70%). The

authors conclude that "to avoid this decline in attitudes toward seeking consent, clerkship directors should ensure that students perform examinations only after patients have given consent explicitly."[22]

Nonexpress forms of consent have been considered and sometimes adopted in different contexts. In late 2006, the US Centers for Disease Control and Prevention (CDC) changed its recommendations about HIV testing and screening for patients in health care settings in which various other diagnostic and screening tests are routinely performed.[23] (Here "diagnostic testing" refers to testing people with clinical signs or symptoms that could indicate HIV infection, while "screening" refers to testing everyone in a certain population.) The policies then in effect, and often embodied in state laws, require specific informed consent, usually in written form, for HIV testing, frequently accompanied by pre-test and post-test counseling. These policies reflected public concerns that had surrounded HIV testing from its beginning in 1985, particularly concerns about the psychosocial risks of stigmatization and discrimination as a result of a positive test. Because of these concerns, testing for HIV was treated differently than testing for other medical conditions, especially those with public health ramifications. Hence, policies at the time required specific disclosure of information and a decision, expressed on a written form, to accept or refuse testing.

The 2006 CDC recommendations moved away from specific written informed consent, accompanied by counseling. In the health care context, the diagnostic testing of patients, in light of clinical signs or symptoms, was justified under implicit consent to medical care, while the screening of all persons ages thirteen to sixty-four, without clinical signs or symptoms of HIV infection, was justified if they were notified that the test would be performed and then given the opportunity to decline. This shift indicated that HIV and AIDS would no longer be treated as exceptions to conventional medical care and to conventional public health measures.[24] The CDC justified its new recommendations primarily on two grounds. First, because HIV and AIDS are chronic conditions that can be effectively treated through anti-retroviral therapies (ARTs), although not cured in the sense of totally and permanently eradicating the virus, the new screening approach would enable more people who are infected to take advantage of available ARTs that could significantly extend their lives at a higher quality. Second, the information gained from screening could enable persons who are infected with HIV to take steps to protect their sex partners or drug-use partners from infection. The CDC estimated that in 2015 over 1.1 million people in the United States were HIV-infected and that one in seven, or approximately 157,000 individuals, were not aware of their infection.[25] Studies after the 2006 recommendations established that treating individuals to reduce their viral load (the concentration of HIV in blood) to undetectable levels can dramatically reduce the risk of spreading HIV infection to sexual or drug-sharing partners.[26] Hence, a slogan arose: "HIV treatment as prevention."[27]

The CDC's changed recommendations did not eliminate patient autonomy in health care settings—individuals could still refuse testing—but, by shifting the default from "opt in" to "opt out," the CDC anticipated that more people previously unaware of their HIV infection would be tested and would gain knowledge that could benefit them and others. Despite these potential benefits, critics warned that in the absence of a requirement for explicit, written informed consent, compromises of autonomy were inevitable in the "opt-out" policy. According to one AIDS activist, "This is not informed consent, and it is not even consent, [but rather an attempt] to ram HIV testing down people's throats without their permission."[28]

In our judgment, this "opt-out" approach, undertaken within CDC guidelines, was and remains justifiable as a way to increase HIV testing without infringing personal autonomy. A strong consensus developed around this approach: By early 2018, all states in the United States had changed their laws regarding HIV testing in medical contexts from "opt-in," through specific, written informed consent, to "opt out."[29]

Another context in which an opt-out approach, sometimes called presumed or tacit consent, could be justified is organ donation from deceased individuals. In the opt-in system in the United States, deceased organ donation requires express, explicit consent, whether by an individual while alive or by the next of kin after his or her death. The information disclosed for the individual's consent is usually limited—for instance, in a cursory exchange when obtaining a license to operate an automobile—but this disclosure is arguably adequate for purposes of postmortem organ donation. In view of the huge gap between the number of organs donated each

year and the number of patients awaiting a transplant, many propose that the United States adopt an opt-out model for organ removal from deceased persons, as several European countries have done. This model shifts the default so that an individual's silence, or nonregistration of dissent, counts as consent, but is such a policy of presumed or tacit consent ethically acceptable?

To be ethically justifiable, such a policy would require vigorous efforts to ensure the public's understanding of the options they face as individuals, as well as a clear, reliable, simple, and nonburdensome mechanism to use to opt out. While accepted in many countries in Europe, an opt-out policy has not yet gained traction in the United States, perhaps because of strong currents of rights of autonomous choice and distrust. Even if it were adopted in the United States, it probably would not increase the number of organs for transplantation overall because, according to survey data, too many citizens would opt out; and opting out would prevent postmortem familial donations, which now provide a large number of transplantable organs when deceased persons have not previously expressed their preferences.[30]

Consents and refusals over time. Beliefs and choices shift over time. Ethical and interpretive problems arise when a person's present choices contradict his or her previous choices, which, in some cases, he or she explicitly designed to prevent possible future changes of mind from affecting an outcome. In one case, a twenty-eight-year-old man decided to terminate chronic renal dialysis because of his restricted lifestyle and the burdens his medical conditions imposed on his family. He had diabetes, was legally blind, and could not walk because of progressive neuropathy. His wife and physician agreed to provide medication to relieve his pain and further agreed not to return him to dialysis even if he requested it under the influence of pain or other bodily changes. (Increased amounts of urea in the blood, which result from kidney failure, can sometimes lead to altered mental states, for example.) While dying in the hospital, the patient awoke complaining of pain and asked to be put back on dialysis. The patient's wife and physician decided to act on the patient's earlier request not to intervene, and he died four hours later.[31]

Their decision was understandable, but respect for autonomy suggests that the spouse and physician should have put the patient back on dialysis to flush the urea out of his bloodstream and then determine if he had autonomously revoked his prior choice. If the patient later indicated that he had not revoked his prior choice, he could have refused again, thereby providing the caregivers with increased assurance about his autonomous preferences.

In shifts over time the key question is whether people are *autonomously* revoking their prior decisions. Discerning whether current decisions are autonomous will depend, in part, on whether they are in character or out of character. Out-of-character actions can raise caution flags that warn others to seek explanations and to probe more deeply into whether the actions are autonomous, but they may turn out to be autonomous. Actions are more likely to be substantially autonomous if they are in character—for example, when a committed Jehovah's Witness refuses a blood transfusion—but acting in character does not necessarily indicate an autonomous action. How, then, are we to determine whether decisions and actions are autonomous?

THE CAPACITY FOR AUTONOMOUS CHOICE

Many patients and potential research subjects are not competent to give a valid consent or refusal. Inquiries about competence focus on whether these persons are capable—cognitively, psychologically, and legally—of adequate decision making. Several commentators distinguish judgments of capacity from judgments of competence on the grounds that health professionals assess capacity and incapacity, whereas courts determine competence and incompetence. However, this distinction breaks down in practice, and we will not rely on it. When clinicians judge that patients lack decision-making capacity, the practical effects of these judgments in a medical context may not differ significantly from those of a legal determination of incompetence.[32]

The Gatekeeping Function of Competence Judgments

Competence or capacity judgments in health care serve a gatekeeping role by distinguishing persons whose decisions should be solicited or accepted from persons whose decisions need not or should not be solicited or accepted. Health professionals' judgments of a person's incompetence may lead them to override that person's decisions, to turn to informal or formal surrogates for decision making, to ask a court to appoint a guardian to protect his or her interests, or to seek that person's involuntary institutionalization. When a court establishes legal incompetence, it appoints a surrogate decision maker with either partial or plenary (full) authority over the incompetent individual.

Competence judgments have the distinctive *normative* function of qualifying or disqualifying persons for certain decisions or actions, but those in control sometimes incorrectly present these competence judgments as *empirical*. For example, a person who appears irrational or unreasonable to others might fail a psychiatric test, and as a result be declared incompetent. The test is an empirical measuring device, but normative judgments establish how the test should be used to sort persons into the two classes of competent and incompetent, which determines how persons ought to be, or may permissibly be, treated.

The Concept of Competence

Some commentators hold that we lack both a single acceptable *definition* of competence and a single acceptable *standard* of competence. They also contend that no nonarbitrary *test* exists to distinguish between competent and incompetent persons. We will engage these issues by distinguishing between definitions, standards, and tests—focusing first on problems of definition.[33]

A single core meaning of the word *competence* applies in all contexts. That meaning is "the ability to perform a task."[34] By contrast to this core meaning, the *criteria* of particular competencies vary from context to context because the criteria are relative to specific tasks. The criteria for someone's competence to stand trial, to raise dachshunds, to answer a physician's questions, and to lecture to medical students are radically different. Rarely should we judge a person as globally incompetent, that is, incompetent with respect to every sphere of life. We usually need to consider only some type of competence, such as the competence to decide about treatment or about participation in research. These judgments of competence and incompetence affect only a limited range of decision making. A person incompetent to decide about financial affairs may be competent to decide whether to participate in medical research.

Competence may vary over time and may be intermittent. Many persons are incompetent to do something at one point in time but competent to perform the same task at another point in time. Judgments of competence about such persons can be complicated by the need to distinguish categories of illness that result in chronic changes of intellect, language, or memory from those characterized by rapid reversibility of these functions, as in the case of transient ischemic attack (TIA) or transient global amnesia (TGA). In some of the latter cases competence varies from hour to hour, and determination of a specific incompetence may prevent vague generalizations that exclude these persons from all forms of decision making.

These conceptual distinctions have practical significance. The law has traditionally presumed that a person incompetent to manage his or her estate is also incompetent to vote, make medical decisions, get married, and the like. The global sweep of these laws, based on a total judgment of the person, at times has extended too far. In a classic case, a physician argued that a patient was incompetent to make decisions because of epilepsy,[35] although many persons who suffer from epilepsy are competent to make decisions in numerous contexts. Such judgments defy much that we now know about the etiology of various forms of incompetence, even in hard cases involving persons with cognitive disabilities, with psychosis, or with uncontrollably painful afflictions. Persons who are incompetent by virtue of dementia, alcoholism, immaturity, or cognitive disabilities present very different types and problems of incompetence.

Sometimes a competent person who ordinarily can select appropriate means to reach his or her goals will act incompetently. Consider the following actual case of a hospitalized patient who has an acute disc problem and whose goal is to control back pain. The patient has decided to manage the problem by wearing a brace, a method she had used successfully in the past. She believes strongly that she should return to this treatment modality.

This approach conflicts, however, with her physician's unwavering and near-insistent advocacy of surgery. When the physician, an eminent surgeon who alone in her city is suited to treat the patient, asks her to sign the surgical permit, she is psychologically unable to refuse. Her illness increases both her hopes and her fears, and, in addition, she has a deferential personality. In these circumstances, it is psychologically too risky for her to act as she prefers. Even though she is competent to choose in general and has stated her preference, she is not competent to choose on this occasion.

This case indicates how close the concept of competence in decision making is to the concept of both autonomy and the principle of respect for autonomy. Patients or prospective subjects are competent to make a decision if they have the capacity to understand the material information, to make a judgment about this information in light of their values, to intend a certain outcome, and to communicate freely their wishes to caregivers or investigators. Although *autonomy* and *competence* differ in meaning (*autonomy* meaning self-governance; *competence* meaning the ability to perform a task or range of tasks), the criteria of the autonomous person and of the competent person are strikingly similar.

Persons are more and less able to perform a specific task to the extent they possess a certain level or range of abilities, just as persons are more and less intelligent or athletic. For example, in the emergency room an experienced and knowledgeable patient is likely to be more qualified to consent to or refuse a procedure than a frightened, inexperienced patient. It would be confusing to view this continuum of abilities in terms of degrees of *competency*. For practical and policy reasons, we need *threshold levels* below which a person with a certain level of abilities for a particular task is incompetent. Where we draw the line depends on the particular tasks involved.[36]

Standards of Competence

Questions in medicine about competence often center on the standards for its determination, that is, the conditions a judgment of competence—and especially incompetence—must satisfy. Standards of competence feature mental skills or capacities closely connected to the attributes of autonomous persons, such as cognitive skills and independent judgment. In criminal law, civil law, and clinical medicine, standards for competence cluster around various abilities to comprehend and process information and to reason about the consequences of one's actions. In medical contexts, physicians often consider a person competent if he or she can understand a procedure, deliberate with regard to its major risks and benefits, and make a decision in light of this deliberation.

The following case illustrates some difficulties encountered in attempts to judge competence. A man who generally exhibits normal behavior patterns is involuntarily committed to a mental institution as the result of the bizarre self-destructive behavior of pulling out an eye and cutting off a hand. This behavior results from his unusual religious beliefs. The institution judges him incompetent, despite his generally competent behavior and despite the fact that his peculiar actions coherently follow from his religious beliefs.[37] This troublesome case is not one of intermittent competence. Analysis in terms of limited competence at first appears plausible, but this analysis perilously suggests that persons with unorthodox or bizarre religious beliefs are less than competent, even if they reason coherently in light of their beliefs. This policy would not be ethically acceptable unless specific and carefully formulated statements spelled out the reasons under which a finding of incompetence is justified.

Rival standards of incompetence. We are focusing on standards of *incompetence*, rather than *competence*, because of the legal, medical, and practical presumption that an adult is competent and should be treated as such in the absence of a determination of incompetence or incapacity. In the clinical context, an inquiry into a patient's competence to make decisions usually occurs only when the medical decision at stake is complex and involves significant risks or when the patient does not accept the physician's recommendation.[38] The following schema expresses the range of inabilities required under *competing standards* of incompetence currently presented in literature on the subject.[39]

1. 1. Inability to express or communicate a preference or choice

2. 2. Inability to understand one's situation and its consequences
3. 3. Inability to understand relevant information
4. 4. Inability to give a reason
5. 5. Inability to give a rational reason (although some supporting reasons may be given)
6. 6. Inability to give risk/benefit-related reasons (although some rational supporting reasons may be given)
7. 7. Inability to reach a reasonable decision (as judged, for example, by a reasonable person standard)

These standards cluster around three kinds of abilities or skills. Standard 1 looks for the ability to formulate a preference, which is an elementary standard. Standards 2 and 3 probe for abilities to understand information and to appreciate one's situation. Standards 4 through 7 concentrate on the ability to reason through a consequential life decision. These standards have been widely used, either alone or in combination, to determine incompetence in medical contexts.

Testing for incompetence. A clinical need exists to turn one or more of these general standards into an operational test of incompetence that establishes passing and failing evaluations. Dementia rating scales, mental status exams, and similar devices test for factors such as time-and-place orientation, memory, understanding, and coherence.[40] Although these clinical assessments are empirical tests, normative judgments underlie each test. The following three ingredients incorporate normative judgments:[41]

1. 1. Choosing the relevant set of abilities for competence
2. 2. Choosing a threshold level of the abilities in item 1
3. 3. Choosing empirical tests for item 2

For any test already accepted under item 3, it is an empirical question whether someone possesses the requisite level of abilities, but this empirical question can only be addressed if normative criteria have already been fixed under items 1 and 2. Institutional rules or traditions usually establish these criteria, but the standards should be open to periodic review and modification.[42]

~~The sliding-scale strategy~~. Some writers offer a sliding-scale strategy for how to realize the goals of competence determinations. They argue that as the risks of a medical intervention increase for patients, so should the level of ability required for a judgment of competence to elect or refuse the intervention. As the consequences for well-being become less substantial, we should lower the level of capacity required for competence. For example, Grisso and Appelbaum present a "competence balance scale." An autonomy cup is suspended from the end of one arm of a measuring scale, and a protection cup is suspended from the other; the fulcrum is set initially to give more weight to the autonomy cup. The balancing judgment depends "on the balance of (1) the patient's mental abilities in the face of the decisional demands, weighed against (2) the probable gain-risk status of the patient's treatment choice."[43] If a serious risk such as death is present, then a correspondingly stringent standard of competence should be used; if a low or insignificant risk is present, then a relaxed or lower standard of competence is permissible. Thus, the same person—a child, for example—might be competent to decide whether to take a tranquilizer but incompetent to decide whether to authorize surgery.[44]

This sliding-scale strategy is attractive. A decision about which standard to use to determine competence depends on several factors that are risk-related. The sliding-scale strategy rightly recognizes that our interests in ensuring good outcomes legitimately contribute to the way we create and apply standards. If the consequences for welfare are grave, the need to certify that the patient possesses the requisite capacities increases; but if little in the way of welfare is at stake, we can lower the level of capacity required for decision making.

Although the sliding-scale strategy may function as a valuable protective device, it creates confusion regarding the nature of both competence judgments and competence itself because of certain conceptual and moral difficulties. This strategy suggests that a person's *competence* to decide is contingent on the decision's importance or on some harm that might follow from the decision. This thesis is dubious: A person's competence to decide whether, for example, to participate in cancer research does not depend on the decision's consequences. As risks increase or decrease, we can legitimately increase or reduce the rules, procedures, or

measures we use to *ascertain* whether someone is competent; but in formulating what we are doing, we need to distinguish between a person's *competence* and the *modes of ascertaining* that person's competence.

Leading proponents of the sliding-scale strategy hold the view that *competence itself* varies with risk. For example, according to Allen Buchanan and Dan Brock, "Because the appropriate level of competence properly required for a particular decision must be adjusted to the consequences of acting on that decision, no single standard of decision-making competence is adequate. Instead, the level of competence appropriately required for decision making varies along a full range from low/minimum to high/maximal."[45]

This account is conceptually and morally perilous. It is correct to say that the level of a person's capacity to decide will rise as the *complexity* or *difficulty* of a task increases (for example, deciding about spinal fusion by contrast to deciding whether to take a minor tranquilizer), but the level of competence to decide does not rise as the *risk* of an outcome increases. It is confusing and misleading to blend a decision's complexity or difficulty with the risk at stake. No basis exists for believing that risky decisions require more ability at decision making than less risky decisions.

We can sidestep these problems by recognizing that the level of *evidence* for determining competence often should vary according to risk. As examples, some statutes have required a higher standard of evidence of competence in making than in revoking advance directives, and the National Bioethics Advisory Commission (NBAC) recommended a higher standard of evidence for determinations of competence to *consent* to participate in most research by contrast to competence to *object* to participation.[46] These are counsels of prudence that stand to protect patient-subjects.

In short, whereas Buchanan and Brock propose that the level of decision-making *competence* itself be placed on a sliding scale from low to high in accordance with risk, we recommend placing the required *standards of evidence* for determining decision-making competence on a sliding scale.

THE MEANING AND JUSTIFICATION OF INFORMED CONSENT

Roughly since the Nuremberg trials, which exposed the Nazis' horrific medical experiments, ethics in medicine and in research has increasingly placed consent at the forefront of its concerns. The term *informed consent* did not appear until a decade after these trials (held in the late 1940s) and did not begin to receive detailed examination until the early 1970s. Over time the physician's or researcher's obligation to *disclose* information shifted significantly to the quality of a patient's or subject's *understanding* and *consent*. The forces behind this shift of emphasis were often autonomy driven. In this section, we treat moral problems of informed consent as they have emerged in clinical ethics, research ethics, case law, changes in the patient-physician relationship, ethics-review committees, and moral and legal theory.[47]

The Justification of Informed Consent Requirements

Virtually all prominent medical and research codes and institutional rules of ethics now state that physicians and investigators must obtain the informed consent of patients and subjects prior to a substantial intervention. Throughout the early history of concern about research subjects, consent requirements were proposed primarily as a way to minimize the potential for harm. However, since the mid-1970s the primary justification of requirements of informed consent has been to protect autonomous choice, a goal that institutions often include in broad statements about protecting the rights of patients and research subjects.

To say that the primary *justification* of informed consent requirements is the protection of autonomy is not to say that the only major *function* of the doctrine and institutions of informed consent is to respect autonomy. As Neal Dickert and coauthors have argued, there may be several distinct functions, including (1) providing transparency; (2) allowing control and authorization; (3) promoting concordance with participants' values; (4) protecting and promoting welfare interests; (5) promoting trust; (6) satisfying regulatory requirements; and (7) promoting integrity in research. These authors hold that "the standard view in research ethics [the "standard

view" being what these authors apparently think is our position] is that the function of informed consent is to respect individual autonomy," which they contend is an unduly narrow conception. We agree that there are multiple functions of informed consent, including their list of seven, although their list of major functions surprisingly omits protection of autonomy. They also judge that in the standard view—presumably our view—there is an "assumption that individual autonomy alone can account for the ethical importance of consent." But we do not hold this view. It is crucial to carefully distinguish *justification* and *function*. Holding that the *justification* of requirements of informed consent is grounded in the principle of respect for autonomy is compatible with recognizing several different *functions* of informed consent requirements.[48]

In a series of books and articles on informed consent and autonomy, Onora O'Neill has argued against the view that informed consent is justified in terms of respect for personal autonomy.[49] O'Neill is suspicious of contemporary conceptions of autonomy and respect for autonomy, which she finds variable, vague, and difficult to tailor to acceptable requirements of informed consent. She argues that practices and rituals of informed consent are best understood as ways to prevent deception and coercion; the process of informed consent provides reasonable assurance that a patient, subject, or tissue donor "has not been deceived or coerced."[50] However, respect for autonomy (and rules of informed consent in health care relationships) requires more than avoiding deception and coercion. It requires an attempt to respect persons' rights to information, improve communication, instill relevant understanding, and avoid forms of manipulation that are not limited to deception and coercion.

The Definition and Elements of Informed Consent

Some commentators have attempted to analyze the idea of informed consent in terms of shared decision making between doctor and patient, thus rendering *informed consent* and *mutual decision making* synonymous.[51] However, informed consent should not be equated with shared decision making. Professionals obtain and will continue to obtain informed consent in many contexts of research and medicine for which shared decision making is a deficient model. We should distinguish (1) informational exchanges and communication processes through which patients and subjects come to elect interventions, often based on medical advice, from (2) acts of approving and authorizing those interventions. Approval and authorization belong to the patient, not to a physician or research investigator, even when extensive shared dialogue has occurred. Shared decision making may appear to be a worthy ideal in some areas of medicine, but the proposed model of sharing decisions is vague and potentially misleading. It cannot be understood as a division of labor, with the clinician deciding A and the patient deciding B. If, alternatively, it is understood as an effort to reach a "joint decision," this position downplays the patient's fundamental ethical and legal right to know and decide.[52] Approving and authorizing are not shared in an appropriate model of informed consent, however much a patient or subject may be influenced by a physician or other health care professionals. In short, this model neither defines nor displaces informed consent; nor does it appropriately implement the principle of respect for autonomy.[53] If shared decision making is presented only as a plea for patients to be *allowed* to participate in decision making about diagnostic and treatment procedures, it continues the legacy of medical paternalism by ignoring patients' *rights* to consent to and authorize or decline those procedures.

Two meanings of "informed consent." Two different senses of "informed consent" appear in current literature, policies, and practices.[54] In the first sense, informed consent is analyzable through the account of autonomous choice presented earlier in this chapter: An informed consent is an individual's *autonomous authorization* of a medical intervention or of participation in research. In this first sense, a person must do more than express agreement or comply with a proposal. He or she must *authorize* something through an act of informed and voluntary consent. In an early and classic case, *Mohr v. Williams* (1905), a physician obtained Anna Mohr's consent to an operation on her right ear. While operating, the surgeon determined that in fact the left ear needed the surgery. A court found that the physician should have obtained the patient's consent to the surgery on the left ear: "If a physician advises a patient to submit to a particular operation, and the patient weighs the dangers and risks incident to its performance, and finally consents, the patient thereby, in effect, enters into a contract authorizing the physician to operate to the extent of the consent given, but no further."[55] An informed consent in

this first sense occurs if and only if a patient or subject, with substantial understanding and in absence of substantial control by others, intentionally authorizes a professional to do something that is specifically mentioned in the consent agreement.

In the second sense, informed consent refers to *conformity to the social rules of consent* that require professionals to obtain legally or institutionally valid consent from patients or subjects before proceeding with diagnostic, therapeutic, or research procedures. Informed consents are not necessarily autonomous acts under these rules and sometimes are not even worded as authorizations. *Informed consent* here refers to an institutionally or legally effective permission, as determined by prevailing social rules. For example, a mature minor may autonomously authorize an intervention, but the minor's authorization may not be an effective consent under existing legal or institutional rules. Thus, a patient or subject might *autonomously* authorize an intervention, and so give an informed consent in the first sense, without *effectively* authorizing the intervention (because of the operative set of rules), and thus without giving an informed consent in the second sense.

Institutional rules of informed consent in law and medicine have frequently not been assessed by the demanding standard of autonomous authorization. As a result, institutions, as well as laws and court decisions, sometimes impose on physicians and hospitals nothing more than an obligation to warn of risks of proposed interventions. "Consent" under these circumstances is not bona fide informed consent in the first sense. The problem arises from the gap between the two senses of informed consent: Physicians who obtain consent under institutional criteria can and often do fail to meet the more rigorous standards of the autonomy-based model.

It is easy to criticize these often lax institutional rules as superficial, but health care professionals cannot reasonably be expected in all circumstances to obtain a consent that satisfies the conditions of highly demanding autonomy-protective rules. Autonomy-protective rules may turn out to be excessively difficult or even impossible to implement in some circumstances. We should evaluate institutional rules in terms of both respect for autonomy and the probable consequences of imposing burdensome requirements on institutions and professionals. Policies may legitimately take account of what is fair and reasonable to require of health care professionals and researchers. Nevertheless, we take as axiomatic that the model of autonomous choice—following the first sense of "informed consent"—ought to serve as the benchmark for the moral adequacy of institutional rules of consent.

Franklin Miller and Alan Wertheimer challenge our view that the first sense of "informed consent" is the benchmark for judging the moral adequacy of institutional understandings and rules of informed consent. They propose a "fair transaction model" of the doctrine of informed consent in which, for example, investigators and their subjects are all treated fairly by giving due consideration to (1) the reasonable limits of an investigator's responsibilities to ensure adequate understanding on the part of subjects who consent to research, (2) the modest levels of comprehension expectable of some subjects, and (3) the overall interests of subjects in participating in research.

We welcome this approach as a reasonable way to think about our second sense of informed consent, but the Miller-Wertheimer theory moves into unacceptably dangerous territory by altogether, and by design, abandoning the first sense of autonomous authorization and substituting the "fair transaction" model. Their model would be more suitable if it were presented as an explication of our second sense of "informed consent" and as a fairness-based analysis of requirements for many practical contexts in which informed consent is obtained. However, as their theory stands, these authors give a priority to fairness to all parties that loses sight of the central role of respect for the subject's or patient's autonomy. We see no justification for their claims that their model merits adoption "in place of the autonomous authorization model" and that "consent is a bilateral transaction," rather than the "one-sided focus on the quality of the subject's consent" to which the autonomous authorization model is committed. Bilateral transactions of informational exchange often appropriately occur in consent contexts, but genuine informed *consent* is not reducible to such transactions.[56]

The elements of informed consent. Some commentators have attempted to define *informed consent* by specifying the essential elements (that is, components) of the concept, in particular by dividing the elements into a set of information components and a set of consent components, and then dividing these components into

subcomponents. The information component refers to the disclosure, and often the comprehension, of information. The consent component refers to both a voluntary decision and an authorization to proceed. Legal, regulatory, philosophical, medical, and psychological literatures generally favor the following elements as the components of informed consent:[57] (1) competence (capacity or ability), (2) disclosure, (3) understanding (comprehension), (4) voluntariness, and (5) consent. Some writers present these elements as the building blocks of a definition of *informed consent* such as the following: A person gives an informed consent to an intervention if (and perhaps only if) he or she is competent to act, receives a thorough disclosure, comprehends the disclosure, acts voluntarily, and consents to the intervention.

This five-element definition is far superior to the single-element definition of *disclosure* that courts and medical literature have often relied on.[58] However, in this chapter we defend and explicate each of the following seven elements as the components of informed consent:

1. I. Threshold elements (preconditions)
    1. 1. Competence (ability to understand and decide)
    2. 2. Voluntariness (in deciding)
2. II. Information elements
    1. 3. Disclosure (of material information)
    2. 4. Recommendation (of a plan)
    3. 5. Understanding (of 3 and 4)
3. III. Consent elements
    1. 6. Decision (in favor of a plan)
    2. 7. Authorization (of the chosen plan)

This list requires explanation. First, *an informed refusal* entails a modification of items under III, thereby turning the categories into refusal elements, for example, "6. Decision (against a plan)." Whenever we use the expression "informed consent," we allow for the possibility of an informed refusal. Second, providing information for potential participants in research does not necessarily involve making a recommendation to the potential participants (number 4), although this component is often the most important from the patient's perspective. Third, competence is perhaps best classified as a *presupposition* of obtaining informed consent, rather than as an *element*.

Having previously examined competence as decision-making capacity, we concentrate in the next three sections on the crucial elements of disclosure, understanding, and voluntariness. These key conditions of informed consent have typically been presumed to be the essential *conceptual* (and perhaps *definitional*) conditions of informed consent, but they can also be viewed as the essential *moral* conditions of a valid consent. As Alexander Capron has appropriately formulated this point, these conditions can be viewed as "the substantive features of [morally] valid informed consent."[59]

DISCLOSURE

Disclosure is the third of the seven elements of informed consent. Some institutions and legal authorities have presented the obligation to disclose information to patients as the sole major condition of informed consent. The legal doctrine of informed consent in the United States from the outset focused primarily, sometimes exclusively, on disclosure because it seemed obvious that physicians must provide sufficient information for a patient to reach a decision and because physicians have an obligation to exercise reasonable care in providing information. Civil litigation has emerged over informed consent because of injuries, measured in terms of monetary damages, that physicians intentionally or negligently have caused by failures to disclose. The term *informed consent* was born in this legal context. However, from the moral viewpoint, informed consent in general has rather little to do with the liability of professionals as agents of disclosure and everything to do with the informed choices of patients and subjects.

Nonetheless, disclosure usually does play a pivotal role in the consent process. Absent professionals' provision of information, many patients and subjects will have an insufficient basis for decision making. Professionals are usually obligated to disclose in reasonably nontechnical language a core body of information, including (1) those facts or descriptions that patients or subjects consider material when deciding whether to refuse or consent to a proposed intervention or involvement in research, (2) information the professional believes to be material, (3) the professional's recommendation (if any), (4) the purpose of seeking consent, and (5) the nature and limits of consent as an act of authorization. If research is involved, disclosures usually should cover the aims and methods of the research, anticipated benefits and risks, any anticipated inconvenience or discomfort, and the subjects' right to withdraw, without penalty, from the research.

This list of basic information could be considerably expanded. For example, in one controversial decision, the California Supreme Court held that, when seeking an informed consent, "a physician must disclose personal interests unrelated to the patient's health, whether research or economic, that may affect the physician's professional judgment."[60] Such a disclosure requirement has acquired increased moral significance as conflicts of interest have become more pronounced and problematic. This subject is examined in Chapter 8.

Standards of Disclosure

Courts have struggled to determine which norms should govern the disclosure of information. Two competing standards of disclosure have become most prominent in the United States: the professional practice standard and the reasonable person standard. A third, the subjective standard, has received some support, although courts have usually avoided it. These standards are morally, not merely legally, important.

The professional practice standard. The first standard holds that a professional community's customary practices determine the adequacy of a disclosure. That is, professional custom establishes the amount and type of information to be disclosed. Disclosure, like treatment, is a responsibility of physicians because of their professional expertise and commitment to the patient's welfare. Accordingly, only expert testimony from members of this profession can count as evidence that a physician violated a patient's right to information.

Several difficulties plague this standard, which some call a reasonable doctor standard because it requires physicians to disclose what any reasonable medical practitioner would disclose in similar cases. First, it is uncertain in many situations whether a customary standard actually exists for the communication of information in medicine. Second, if custom alone were conclusive, pervasive negligence could be perpetuated with impunity. The majority of professionals could offer the same inadequate level of information. Third, based on empirical studies, it is questionable whether many physicians have developed the skills to determine the information that serves their patients' best interests.[61] The weighing of risks in the context of a person's subjective beliefs, fears, and hopes is not an expert skill, and information provided to patients and subjects sometimes needs to be freed from the entrenched values and goals of medical professionals. Finally, the professional practice standard ignores and may subvert patients' rights of autonomous choice. Professional standards in medicine are fashioned for medical judgments, but final decisions for or against medical interventions are nonmedical decisions that belong solely to the patient.

The reasonable person standard. Despite the adoption of the traditional professional practice standard in many legal jurisdictions, a reasonable person standard has gained acceptance in many states in the United States. According to this standard, the information to be disclosed should be determined by reference to a hypothetical reasonable person. Whether information is pertinent or material is to be measured by the significance a reasonable person would attach to it in deciding whether to undergo a procedure. Under this standard the authoritative determination of informational needs shifts from the physician to the patient, and physicians may be found guilty of negligent disclosures even if their behavior conforms to recognized professional practice.

Whatever its merits, the reasonable person standard presents conceptual, moral, and practical difficulties. Unclarities surround the concepts of "material information" and "reasonable person," and questions arise about whether and how physicians and other health care professionals can employ the reasonable person standard in

practice. Its abstract and hypothetical character makes it difficult for physicians to use because they must project by hypothesis what a reasonable patient would need to know.

The subjective standard. The reasonable person standard is widely considered to be an *objective standard*. By contrast, the *subjective standard* judges the adequacy of information by reference to the specific informational needs of the individual person rather than by the hypothetical reasonable person. Individual needs can differ: Persons may have unconventional beliefs, unusual health problems, or unique family histories that require a different informational base than the objective reasonable person needs. For example, a person with a family history of reproductive problems might desire information that other persons would not need or want before becoming involved in research on sexual and familial relations. If a physician knows or has reason to believe that a person wants such information, then withholding it may undermine autonomous choice. The key issue is whether the standard for the disclosure of information should be tailored to the individual patient and thus made subjective.[62]

Of the three standards, the subjective standard is the preferable *moral* standard of disclosure, because it alone takes the idea of respect for autonomy seriously and meets persons' specific informational needs. Nevertheless, an exclusive reliance on the subjective standard would not suffice for either law or ethics because patients often do not know what information is relevant for their deliberations, and we cannot reasonably expect a doctor to do an exhaustive background and character analysis of each patient to determine the relevant information. Hence, for purposes of ethics, it is best to use the reasonable person standard as the initial standard of disclosure and then supplement it by investigating the informational needs of particular patients or potential research subjects.

Intentional Nondisclosure

Numerous topics in bioethics involve problems of intentional nondisclosure. They include medical confidentiality, informed refusal, placebo treatment, randomized clinical trials, genetic counseling, and the duty to warn third parties. In each area questions have arisen about whether withholding information to patients or subjects is justified and, if so, under which conditions. For example, in randomized clinical trials, patients commonly do not know whether they are receiving an investigational drug of interest or rather are receiving no treatment at all. Some argue that it is ethically acceptable, and highly desirable in some situations, to randomize patients without their express knowledge and consent in trials comparing widely used, approved interventions that pose no additional risk.[63] However, ethical controversies have erupted over failures to obtain adequately informed consent for some clinical trials comparing different accepted treatments; a primary example is the SUPPORT study of oxygen therapy for premature babies.[64]

In this section we begin with two problems of intentional nondisclosure in clinical ethics and then turn to problems of withholding information from research subjects. All three subsections ask, "Are these intentional nondisclosures justifiable?"

Therapeutic privilege. Several controversies in clinical practice involve questions about the conditions under which a person's right to autonomous choice demands a disclosure by a physician that would either harm the patient or harm someone connected to the patient such as a family member or partner. As contexts change—for example, as a patient becomes increasingly frightened or agitated—the weights of competing moral demands of respect for autonomy and beneficence vary, and no decision rule is available to determine whether and when one obligation outweighs the other. No one in bioethics has formulated a hierarchical-ordering rule that requires that respect for the autonomy of patients and full disclosure of information always overrides the physician's obligations to make a good medical judgment about how to protect patients from harm-causing conditions, and no general theoretical considerations show that physicians must never intentionally withhold information. Much depends on the weight, in any given circumstance, of a medical benefit and the importance of an item of information for the patient. (This general problem is explored in the section of Chapter 6 entitled "Paternalism: Conflicts between Beneficence and Respect for Autonomy" and in the discussion of "Veracity" in Chapter 8.)

Legal exceptions to the rule of informed consent often allow a health professional to proceed without consent in cases of emergency, incompetence, and waiver. The first two exceptive conditions are generally uncontroversial,

but some controversy surrounds waivers. A notably controversial exception is the therapeutic privilege, which states that a physician may legitimately withhold information based on a sound medical judgment that divulging the information would potentially harm a depressed, emotionally drained, or unstable patient. Possible harmful outcomes include endangering life, causing irrational decisions, and producing anxiety or stress.[65]

Despite this exception's traditionally protected status, United States Supreme Court Justice Byron White once vigorously attacked the idea that the possibility of increasing a patient's anxiety about a procedure provides sufficient justification for an exception to rules of informed consent. White suggested that the legally protected status of the doctrine of therapeutic privilege lacks the security it once had in medicine.[66]

Attempts to justify the therapeutic privilege are beneficence- and nonmaleficence-based because nondisclosure is aimed at the patient's welfare and at preventing harm from occurring. However, the precise content and formulation of the therapeutic privilege varies across legal jurisdictions and institutional practices. Some formulations permit physicians to withhold information if disclosure would cause *any* deterioration in the patient's condition. Other formulations permit the physician to withhold information if and only if the patient's knowledge of the information would have serious health-related consequences such as jeopardizing the treatment's success or critically impairing the patient's relevant decision-making faculties.

The narrowest formulation of the therapeutic privilege appeals to circumstances of incompetence: A physician may invoke the therapeutic privilege only if he or she has sufficient reason to believe that disclosure would render the patient incompetent to consent to or refuse the treatment. This criterion does not conflict with respect for autonomy, because the patient would be incapable of an autonomous decision at the point the decision would occur. However, it is ethically indefensible, even if legally permissible, to invoke the therapeutic privilege merely on grounds that the disclosure of relevant information might lead a competent patient to refuse a proposed treatment.[67]

Therapeutic use of placebos. A related problem in clinical ethics is the *therapeutic use of placebos*, which typically, but not always or necessarily, involves limited transparency, incomplete disclosure, or even intentional deception. A placebo is a substance or intervention that the clinician believes to be pharmacologically or biomedically inert or inactive for the condition being treated. While "pure" placebos, such as a sugar pill, are pharmacologically inactive, active medications are sometimes used as "impure" placebos for conditions for which they are not medically indicated—for example, the prescription of an antibiotic for a common cold. Systematic evidence is lacking for the clinically significant benefits of most placebos,[68] but patient and clinician reports indicate that placebos relieve some subjective symptoms in as many as one-third of patients who suffer from conditions such as angina pectoris, cough, anxiety, depression, hypertension, headache, and the common cold.[69] Placebos have also been reported to help some patients with irritable bowel syndrome, pain, and nausea.[70] The primary benefits of placebos occur for more subjective and self-reported symptoms, that is, for the illness as experienced, rather than for the underlying disease. For instance, a small study of patients with asthma compared active albuterol, which is a standard treatment, with placebo, sham acupuncture, and no intervention.[71] Only active albuterol improved forced expiratory volume (FEV), an important measure of pulmonary function. However, according to self-reported outcomes, active albuterol provided no incremental benefit over placebo and sham acupuncture. While acknowledging such subjective self-reports, critics focus on placebos' lack of effect on underlying diseases.

Despite the limited evidence for the clinical benefits of placebos, their provision or prescription is common in clinical practice. In a national study of US internists and rheumatologists, approximately half of the respondents reported that over the previous year they had prescribed placebo treatments on a regular basis, most often over-the-counter analgesics and vitamins. Slightly more than 10% had prescribed antibiotics or sedatives as placebo treatments; only a few had used saline or sugar pills as placebo treatments. Over 60% of those surveyed expressed a belief that the practice of prescribing placebos is ethically permissible.[72] One survey of patients with a chronic health problem for which they had seen a primary care provider at least once over the previous six months found that most were receptive to physicians' provision or prescription of placebo treatments, depending

on the circumstances, especially the conditions of transparency and honesty. Only 21.9% opposed placebo treatments under any circumstances.[73]

Beyond arguments against deception and failure to respect autonomy, objections to the therapeutic provision or prescription of placebos include possible negative consequences such as damage to a specific clinical relationship or to clinical relationships in general because of reduced trust.[74] Some defenses of placebos hold that a patient's consent to a generic treatment such as "an effective pill" or "a powerful medicine" is sufficient. A related defense of placebos appeals to the patient's prior consent to the goals of treatment. Such consent is not informed consent, but these proposals could be rendered acceptable if the patient were informed in advance that a placebo would or might be used at some point in the course of treatment and the patient consented to this arrangement.[75]

Taking a somewhat similar approach, the American Medical Association (AMA) updated its policy on the therapeutic use of placebos in 2016. It set three necessary conditions for a physician to meet before using a placebo for diagnosis or treatment: (1) enlist the cooperation of the patient, (2) obtain the patient's "general consent to administer a placebo," and (3) avoid using a placebo merely to manage a difficult patient. By obtaining "general consent" (the second condition), the physician "respects the patient's autonomy and fosters a trusting relationship while the patient still may benefit from the placebo effect."[76]

Evidence indicates that the placebo response or placebo effect can sometimes be produced without nondisclosure or deception. For example, the placebo response or effect sometimes occurs even if patients have been informed that a particular substance is pharmacologically inert and consent to its use.[77] The mechanisms of placebo responses are poorly understood, but several hypotheses have been proposed, frequently centering on the healing context and its symbolic significance and rituals (including the ritual of taking medications) and on the professional's care, compassion, and skill in fostering trust and hope.[78] However, it is important, when prescribing placebos, that clinicians not bypass opportunities for effective communication with patients. Effective communication and enhanced patient understanding can be fostered by admitting uncertainty; exploring patients' concerns, outlooks, and values; and inviting patients to be partners in the search for therapeutic options.[79]

Withholding information from research subjects. Problems of intentional nondisclosure in clinical practice have parallels in forms of research in which investigators sometimes withhold information from subjects. Occasionally, good reasons support nondisclosure. For instance, scientists could not conduct vital research in fields such as epidemiology if they always had to obtain consent from subjects for access to their medical records. They justify using such records without consent to establish the prevalence of a particular disease. This research is often the first phase of an investigation intended to determine whether to trace and contact particular individuals who are at risk of disease, and the researchers may need to obtain their permission for further participation in research. Sometimes, however, researchers are not required to contact individuals at all, as, for example, when hospitals strip personal identifiers from their records so that epidemiologists cannot identify individual patients. In other circumstances, researchers only need to notify persons in advance about how they will use data and offer them the opportunity to refuse to participate. In short, some disclosures, warnings, and opportunities to decline involvement are legitimately substituted for informed consent.

Other forms of intentional nondisclosure in research are difficult to justify. For instance, vigorous debate arose about a study, designed and conducted by two physicians at the Emory University School of Medicine, to determine the prevalence of cocaine use and the reliability of self-reports of drug use among male patients in an Atlanta walk-in, inner-city hospital clinic serving low-income, predominantly black residents. In this study, approved by the institutional human investigations committee, researchers asked weekday outpatients at Grady Memorial Hospital to participate in a study about asymptomatic carriage of sexually transmitted diseases (STDs). The participants provided informed consent for the STD study, but not for an *unmentioned* piggy-back study on recent cocaine use and the reliability of self-reports of such use. Researchers informed patients that their urine would be tested for STDs, but did not disclose that their urine would also be tested for cocaine metabolites. Of the 415 eligible men who agreed to participate, 39% tested positive for a major cocaine

metabolite, although 72% of those with positive urinary assays denied any illicit drug use in the three days prior to sampling. Researchers concluded: "Our findings underscore the magnitude of the cocaine abuse problem for young men seeking care in inner-city, walk-in clinics. Health care providers need to be aware of the unreliability of patient self-reports of illicit drug use."[80]

This finding was valuable at the time, but these researchers deceived their subjects about some aims and purposes of the research and did not disclose the means they would use. Investigators thought they faced a dilemma: On the one hand, they needed accurate information about illicit drug use for health care and public policy. On the other hand, obtaining adequate informed consent would be difficult, because many potential subjects would either refuse to participate or would offer false information to researchers. The moral problem is that rules requiring informed consent have been designed to protect subjects from manipulation and abuse during the research process. Reports of the strategy used in this cocaine study could increase suspicion of medical institutions and professionals and could make patients' self-reports of illegal activities even less reliable.[81] Investigators could have resolved their dilemma by developing alternative research designs, including sophisticated methods of using questions that can either reduce or eliminate response errors without violating rules of informed consent.

In general, research cannot be justified if significant risk is involved and subjects are not informed that they are being placed at risk. This conclusion does not imply that researchers can never justifiably undertake studies involving deception. Relatively risk-free research involving deception or incomplete disclosure has been common in fields such as behavioral and physiological psychology. However, researchers should use deception only if it is essential to obtain vital information, it involves no substantial risk to subjects and society, subjects are informed that deception or incomplete disclosure is part of the study, and subjects consent to participate under these conditions. (Similar problems of research ethics are discussed in Chapter 8 in the sections on "Veracity" and "The Dual Roles of Clinician and Investigator.")

UNDERSTANDING

Understanding is the fifth element of informed consent in our earlier list. Clinical experience and empirical data indicate that patients and research subjects exhibit wide variation in their understanding of information about diagnoses, procedures, risks, probable benefits, and prognoses.[82] In a study of participants in cancer clinical trials, 90% indicated they were satisfied with the informed consent process and most thought they were well informed. However, approximately three-fourths of them did not understand that the trials included nonstandard and unproven treatment, and approximately one-fourth did not appreciate that the primary purpose of the trials was to benefit future patients and that the benefits to them personally were uncertain.[83]

Many factors account for limited understanding in the informed consent process. Some patients and subjects are calm, attentive, and eager for dialogue, whereas others are nervous or distracted in ways that impair or block understanding. Illness, irrationality, and immaturity also can limit understanding. Important institutional and situational factors include pressures of time, limited or no remuneration to professionals for time spent in communication, and professional conflicts of interest.

The Nature of Understanding

No general consensus exists about the nature and level of understanding needed for an informed consent, but an analysis sufficient for our purposes is that persons understand if they have acquired pertinent information and have relevant beliefs about the nature and consequences of their actions. Their understanding need not be *complete*, because a grasp of central facts is usually sufficient. Some facts are irrelevant or trivial; others are vital, perhaps decisive.

In some cases, a person's lack of awareness of even a single risk or missing fact can deprive him or her of adequate understanding. Consider, for example, the classic case of *Bang v. Miller Hospital* (1958), in which

patient Helmer Bang suffered from urinary problems for which he sought treatment, but he did not intend to consent to a sterilization entailed in the recommended prostate surgery.[84] Bang did, in fact, consent to prostate surgery, but without being told that sterilization was an inevitable outcome. Although sterilization is not necessarily an outcome of prostate surgery, it was inevitable in the specific procedure recommended, which involved cutting Bang's spermatic cord. Bang's failure to understand this one surgical consequence compromised what was otherwise an adequate understanding and invalidated what otherwise would have been a valid consent.

Patients and subjects usually should understand, at a minimum, what an attentive health care professional or researcher believes a reasonable patient or subject needs to understand to authorize an intervention. Diagnoses, prognoses, the nature and purpose of the intervention, alternatives, risks and benefits, and recommendations typically are essential. Patients or subjects also need to share an understanding with professionals about the terms of the authorization before proceeding. Unless agreement exists about the essential features of what is authorized, there is no assurance that a patient or subject has made an autonomous decision and provided a valid consent. Even if the physician and the patient both use a word such as *stroke* or *hernia*, their interpretations may diverge if standard medical conceptions as used by the physician have meanings the patient does not understand.

Some argue that many patients and subjects cannot comprehend enough information or sufficiently appreciate its relevance to make autonomous decisions about medical care or participation in research. Such statements overgeneralize, often because of an improper ideal of full disclosure and full understanding. If we replace this unrealistic standard with a more defensible account of the understanding of material information, we can avoid this skepticism. From the fact that actions are never *fully* informed, voluntary, or autonomous, it does not follow that they are never *adequately* informed, voluntary, or autonomous.[85]

However, some patients have such limited knowledge bases that communication about alien or novel situations is exceedingly difficult, especially if physicians introduce new concepts and cognitive constructs. Various studies indicate that these patients likely will have an impoverished and distorted understanding of scientific goals and procedures.[86] However, even in these difficult situations enhanced understanding and adequate decision making can often be achieved. Professionals may be able to communicate novel or specialized information to laypersons by drawing analogies between this information and more ordinary events familiar to the patient or subject. Similarly, professionals can express risks in both numeric and nonnumeric probabilities, while helping the patient or subject to assign meanings to the probabilities through comparison with more familiar risks and prior experiences, such as risks involved in driving automobiles or using power tools.[87]

Even with the assistance of such strategies, enabling a patient to both comprehend and appreciate risks and probable benefits can be a formidable task. For example, patients confronted with various forms of surgery understand that they will suffer post-operative pain, but their projected expectations of pain are often inadequate. Many patients cannot in advance adequately appreciate the nature and severity of the pain, and many ill patients reach a point when they can no longer balance with clear judgment the threat of pain against the benefits of surgery. At this point, they may find the benefits of surgery overwhelmingly attractive, while discounting the risks.

Studies of comprehension. Some studies focus on patients' and research participants' failures to comprehend the risks involved, but problems also arise in the understanding of expected benefits—their nature, probability, and magnitude. These problems were evident in a study of the understanding of patients with stable coronary artery disease who chose to undergo percutaneous coronary intervention (PCI). In contrast to the best available evidence and the views of their cardiologists, the overwhelming majority of these patients thought that PCI would reduce their risk of a heart attack (88%) and their risk of death from a heart attack (82%), even though PCI's major expected benefit for such patients is only symptomatic, namely, relief from chest pain or discomfort. PCI may be lifesaving for patients who have an acute or unstable angina, and the patients who had only stable angina may have confused the two conditions because both involve chest pain and discomfort. According to the investigators and a commentator, direct communication about these and other matters, accompanied by decision

aids, could have been helpful, especially when accompanied by improvements in the level of reading difficulty and the information provided in the consent form.[88]

The therapeutic misconception. The "therapeutic misconception" is an important problem of informed consent that must be addressed where subjects may fail to distinguish between clinical care and nontherapeutic research and may fail to understand the purpose and aim of research, thereby misconceiving their participation as therapeutic in nature.[89] The therapeutic misconception presumably invalidates a subject's consent because he or she is not specifically consenting to participation *in research*.[90]

Sam Horng and Christine Grady appropriately distinguish therapeutic misconception in the strict sense from therapeutic misestimation and therapeutic optimism.[91] The therapeutic misconception, if uncorrected, invalidates subjects' consent because they do not have relevant facts sufficiently straight to consent to participate in research. However, some participants who understand that they are involved in research rather than clinical care still overestimate the therapeutic possibilities and probabilities—that is, the odds that participants will benefit. Such a therapeutic misestimation, Horng and Grady argue, should be tolerated if "modest misestimates do not compromise a reasonable awareness of possible outcomes." By contrast, in therapeutic optimism participants accurately understand the odds that participants will benefit but are overly optimistic about their own chances of beating those odds. This therapeutic optimism usually does not compromise or invalidate the individual's informed consent because it more approximates a legitimate hope than an informational bias.

Problems of Information Processing

With the exception of a few studies of comprehension, studies of patients' decision making pay insufficient attention to information processing. Yet information overload may prevent adequate understanding, and physicians exacerbate these problems when they use unfamiliar medical terms.

Some studies have uncovered difficulties in processing information about risks, indicating that risk disclosures commonly lead subjects to distort information, promote inferential errors, and create disproportionate fears of some risks. Some ways of framing information are so misleading that both health professionals and patients regularly misconstrue the content. For example, choices between risky alternatives can be influenced by whether the same risk information is presented as providing a gain or an opportunity for a patient or as constituting a loss or a reduction of opportunity.[92]

One study asked radiologists, outpatients with chronic medical problems, and graduate business students to make a hypothetical choice between two alternative therapies for lung cancer: surgery and radiation therapy.[93] Researchers framed the information about outcomes in terms of (1) survival and (2) death. This difference of framing affected preferences in all three groups. When faced with outcomes framed in terms of probability of *survival*, 25% chose radiation over surgery. However, when the identical outcomes were presented in terms of probability of *death*, 42% preferred radiation. The mode of presenting the risk of immediate death from surgical complications, which has no counterpart in radiation therapy, appears to have made the decisive difference.

These framing effects reduce understanding, with direct implications for autonomous choice. If a misperception prevents a person from adequately understanding the risk of death and this risk is material to the person's decision, then the person's choice of a procedure does not reflect a substantial understanding and his or her consent does not qualify as an autonomous authorization. The lesson is that professionals need greater knowledge of techniques that can enable them to communicate better both the positive and the negative facets of information—for example, both the survival and the mortality probabilities.

Decision aids are increasingly used to prepare individuals to participate in medical decisions that involve balancing probable benefits and risks in contexts of scientific uncertainty where decisions about screening or therapeutic interventions are difficult to evaluate. Studies show that the use of decision aids can provide important information and enable patients to reflect on their own values and preferences in relation to their

circumstances and options. The use of these decision aids correlates with patients' increased knowledge and more active participation in decision making.[94]

Problems of Nonacceptance and False Belief

A breakdown in a person's ability to *accept* information as true or untainted, even if he or she adequately comprehends the information, also can compromise decision making. A single false belief can in some circumstances invalidate a patient's or subject's consent, even when there has been a suitable disclosure, comprehension, and voluntary decision making by the patient. For example, a seriously ill patient who has been adequately informed about the nature of the illness and has been asked to make a treatment decision might refuse under the false belief that he or she is not ill. Even if the physician recognizes the patient's false belief and adduces conclusive evidence to prove to the patient that the belief is mistaken, and the patient comprehends the information provided, the patient may go on believing that what has been reported is false.

If ignorance prevents an informed choice, it may be permissible and possibly obligatory to promote autonomy by attempting to impose unwelcome information. Consider the following case in which a false belief played a major role in a patient's refusal of treatment:[95]

> A fifty-seven-year-old woman was admitted to the hospital because of a fractured hip. … During the course of the hospitalization, a Papanicolaou test and biopsy revealed stage 1A carcinoma of the cervix. … Surgery was strongly recommended, since the cancer was almost certainly curable by a hysterectomy. … The patient refused the procedure. The patient's treating physicians at this point felt that she was mentally incompetent. Psychiatric and neurological consultations were requested to determine the possibility of dementia and/or mental incompetency. The psychiatric consultant felt that the patient was demented and not mentally competent to make decisions regarding her own care. This determination was based in large measure on the patient's steadfast "unreasonable" refusal to undergo surgery. The neurologist disagreed, finding no evidence of dementia. On questioning, the patient stated that she was refusing the hysterectomy because she *did not believe* she had cancer. "Anyone knows," she said, "that people with cancer are sick, feel bad and lose weight," while she felt quite well. The patient continued to hold this view despite the results of the biopsy and her physicians' persistent arguments to the contrary.

The physician in this case considered overriding the patient's refusal, because solid medical evidence indicated that she was unjustified in believing that she did not have cancer. As long as this patient continues to hold a false belief that is material to her decision, her refusal is not an adequately *informed* refusal even if it might turn out to be a legally valid refusal. The case illustrates some complexities involved in effective communication: The patient was a poor white woman from Appalachia with a third-grade education. The fact that her treating physician was black was the major reason for her false belief that she did not have cancer. She would not believe what a black physician told her. However, intense and sometimes difficult discussions with a white physician and with her daughter eventually corrected her belief and led her to consent to a successful hysterectomy.

This example illustrates why it is sometimes necessary for clinicians to vigorously challenge patients' choices that appear to be legally binding in order to further enhance the quality of their choices rather than merely accept their choices at face value. The right to refuse unwanted treatment has the appearance of a near absolute right in biomedical ethics, but the case just considered indicates that health care professionals should carefully consider when this right needs to be challenged and perhaps even overridden.

Problems of Waivers

Further problems about understanding arise in waivers of informed consent. In the exercise of a waiver, a competent patient voluntarily relinquishes the right to an informed consent and relieves the physician of the obligation to obtain informed consent.[96] The patient delegates decision-making authority to the physician or to a third party, or simply asks not to be informed; the patient in effect makes a decision not to make an informed

decision. However, waivers need not be understood exclusively in this way. Regulations recognize various waivers of consent requirements as valid when patients or subjects do not autonomously authorize by a waiver in the normal sense. Examples of such valid waivers occur under conditions of impracticability, emergency research, and drug and vaccine research with armed forces personnel.[97]

Some courts have held that physicians need not make disclosures of risk if a patient requests not to be informed,[98] and some writers in biomedical ethics hold that rights are always waivable.[99] It is usually appropriate to recognize waivers of rights because we enjoy discretion over whether to exercise such rights. For example, if a committed Jehovah's Witness informed a doctor that he wished to have everything possible done for him but did not want to know if the hospital utilized transfusions or similar procedures, it is difficult to imagine a moral argument sufficient to support the conclusion that he must give a specific informed consent to the transfusions. Nevertheless, a general practice of allowing waivers is dangerous. Many patients have an inordinate trust in physicians, and a widespread acceptance of waivers of consent in research and therapeutic settings could make subjects and patients more vulnerable to those who omit consent procedures for convenience, which is already a serious problem in health care.

No solution to these problems about waivers is likely to emerge that fits all cases. Although each case or situation of waiver needs to be considered separately, appropriate procedural responses that provide oversight to protect patients may be needed. For example, institutions can develop rules that disallow waivers except when they have been approved by deliberative bodies, such as institutional review committees and hospital ethics committees. If a committee determines that recognizing a waiver would best protect a person's interest in a particular case, the waiver could justifiably be sustained.

VOLUNTARINESS

Voluntariness is another element of informed consent and also the third of our three conditions of autonomous action. Because it was often neglected in the history of research, this element has come to have a prominent role in biomedical ethics. The Nuremberg Code, for example, insists on voluntariness: A research subject "should be so situated as to be able to exercise free power of choice, without the intervention of any element of force, fraud, deceit, duress, over-reaching, or other ulterior form of constraint or coercion."[100]

We use the term *voluntariness* more narrowly than some writers do. Some have analyzed voluntariness in terms of the presence of adequate knowledge, the absence of psychological compulsion, and the absence of external constraints.[101] If we were to adopt such a broad meaning, we would be equating voluntariness with autonomy, whereas our claim is only that voluntariness—here understood primarily as freedom from controlling conditions—is a necessary condition of autonomy. A person acts voluntarily if he or she wills the action without being under the control of another person or the control of a personal psychological condition. We consider here only the condition of control by other individuals, but we note that conditions such as debilitating disease, psychiatric disorders, and drug addiction can diminish or destroy voluntariness, thereby precluding autonomous choice and action.

Forms of Influence

Not being controlled is the key condition of voluntariness, but not all influences exerted on another person are controlling. If a physician orders a reluctant patient to undergo cardiac catheterization and coerces the patient into compliance through a threat of abandonment, then the physician's influence controls the patient. If, by contrast, a physician rationally persuades the patient to undergo the procedure when the patient is at first reluctant to do so, then the physician's actions influence but do not control the patient. Many influences are resistible, and some are welcomed rather than resisted.

The broad category of influence includes acts of love, threats, education, lies, manipulative suggestions, and emotional appeals, all of which can vary dramatically in their impact on persons and in their ethical justification.

Our analysis focuses on three categories of influence: coercion, persuasion, and manipulation. Coercion occurs if and only if one person intentionally uses a credible and severe threat of harm or force to control another.[102] The threat of force used by some police, courts, and hospitals in acts of involuntary commitment for psychiatric treatment is coercive. Some threats will coerce virtually all persons (e.g., a credible threat to kill the person), whereas others will coerce only a few persons (e.g., an employee's threat to an employer to quit a job unless a raise is offered). Whether coercion occurs depends in part on the subjective responses of the coercion's intended target. However, a subjective response in which persons comply because they *feel* threatened even though no threat has actually been issued does not qualify as coercion. Coercion occurs only if an intended and credible threat displaces a person's self-directed course of action, thereby rendering even intentional and well-informed behavior nonautonomous. We reject a common tendency in biomedical ethics to use "coercion" as a broad term of ethical criticism that obscures relevant and distinctive ethical concerns. For instance, coercion is not identical to taking advantage of a person in dire circumstances. Both are wrong in most contexts, but perhaps for different reasons.[103]

In *persuasion* a person comes to believe something through the merit of reasons another person advances. Appeal to reason is distinguishable from influence by appeal to emotion. In health care, the problem is how to distinguish emotional responses from cognitive responses and to determine which are likely to be evoked. Disclosures or approaches that might rationally persuade one patient might overwhelm another whose fear or panic undercuts reason.

Manipulation is a generic term for several forms of influence that are neither persuasive nor coercive.[104] The essence of manipulation is swaying people to do what the manipulator wants by means other than coercion or persuasion. In health care the most common form of manipulation is informational manipulation, a deliberate act of managing information that alters a person's understanding of a situation and motivates him or her to do what the agent of influence intends. Many forms of informational manipulation are incompatible with autonomous decision making. For example, lying, withholding information, and exaggeration with the intent to lead persons to believe what is false all compromise autonomous choice. The manner in which a health care professional presents information—by tone of voice, by forceful gesture, and by framing information positively ("we succeed most of the time with this therapy") rather than negatively ("we fail with this therapy in 35% of the cases")—can also manipulate a patient's perception and response.

Nevertheless, it is easy to inflate control by manipulation beyond its actual significance in health care. We often make decisions in a context of competing influences, such as personal desires, familial constraints, legal obligations, and institutional pressures, but these influences usually do not control decisions to a morally worrisome degree.

The Obligation to Abstain from Controlling Influence

Coercion and controlling manipulation are occasionally justified—infrequently in medicine, more often in public health, and even more often in law enforcement. If a physician taking care of a disruptive and noncompliant patient threatens to discontinue treatment unless the patient alters certain behaviors, the physician's mandate may be both coercive and justified. The most difficult problems about manipulation do not involve threat and punishment, which are almost always unjustified in health care and research. They involve the effect of rewards, offers, encouragement, and other nudges.

A classic example of an unjustified offer occurred during the Tuskegee syphilis study, which left close to four hundred African American males who had been diagnosed with syphilis untreated for decades in order to study the natural history of untreated syphilis, even though penicillin, an effective treatment for syphilis, became available during those years. Researchers used various offers to stimulate and sustain the subjects' interest in continued participation; these offers included free burial assistance and insurance, free transportation to and from the examinations, and a free stop in town on the return trip. Subjects also received free medicines and free hot meals on the days of their examination. The subjects' socioeconomic deprivation made them vulnerable to

these overt and unjustified forms of manipulation.[105] These manipulative endeavors were coupled with deception that hid the nature and nontherapeutic intent of the study.

The conditions under which an influence both controls persons and lacks moral justification are reasonably clear in theory but often unclear in concrete situations. For example, patients have reported feeling severe pressure to enroll in clinical trials, even though their enrollment is voluntary.[106] Some difficult cases in health care involve manipulation-like situations in which patients or subjects desperately need a given medication or a source of income. Attractive offers such as free medication or extra money can leave a person without a meaningful choice. A threatening situation can constrain a person even in the absence of another's intentional manipulation. Influences that persons ordinarily find resistible can control abnormally weak, dependent, and surrender-prone patients.[107] People's vulnerabilities differ, producing variations in what constitutes an "undue" influence.[108]

The threat of exploitation for research and other purposes is substantial in institutions in which populations are confined involuntarily. Rules, policies, and practices can work to compromise autonomous choice even if persons voluntarily admit themselves to institutions. Consider long-term care, where the elderly in nursing homes can experience constricted choices in everyday matters. Many suffer a decline in the ability to carry out personal choices because of physical impairments, but this decline in *executional* autonomy need not be accompanied by a decline in *decisional* autonomy.[109] On the one hand, the problem is that caregivers in nursing homes may neglect, misunderstand, or override residents' autonomous decisions in everyday decisions about food, roommates, possessions, exercise, sleep, and clothes, along with baths, medications, and restraints. On the other hand, institutional needs for structure, order, safety, and efficiency are sometimes legitimately invoked to override residents' apparent autonomous choices.

SURROGATE DECISION MAKING FOR ECISION NONAUTONOMOUS PATIENTS

We turn now from conditions of consent by autonomous decision makers—and limitations on autonomy in some situations—to standards of surrogate decision making when patients are not autonomous or are doubtfully autonomous. Surrogates daily make decisions to terminate or continue treatment for incompetent patients, for example, those suffering from stroke, Alzheimer's disease, Parkinson's disease, chronic depression affecting cognitive function, senility, and psychosis. If a patient is not competent to accept or refuse treatment, a hospital, physician, or family member may justifiably exercise a decision-making role, depending on legal and institutional rules, or go before a court or other authority to resolve uncertainties about decision-making authority.

Three general standards have been proposed for use by surrogate decision makers: *substituted judgment*, which is sometimes presented as an autonomy-based standard; *pure autonomy*; and *the patient's best interests*. Our objective in this section is to restructure and integrate this set of standards for surrogate decision, creating a coherent framework. We evaluate these standards for purposes of law and policy, but our underlying moral argument concerns how to protect both patients' former autonomous preferences and their current best interests. (In Chapter 5 we examine *who* should be the surrogate decision maker.)

The Substituted Judgment Standard

The standard of substituted judgment holds that decisions about treatment properly belong to the incompetent or nonautonomous patient because of his or her rights of autonomy and privacy. Patients have the right to decide and to have their values and preferences taken seriously even when they lack the capacity to exercise those rights. It would be unfair to deprive an incompetent patient of decision-making rights merely because he or she is no longer, or has never been, autonomous.

This is a weak standard of autonomy. It requires the surrogate decision maker to "don the mental mantle of the incompetent," as a judge in a classic court case put it; the surrogate is to make the decision the incompetent

person would have made if competent. In this case, the court invoked the standard of substituted judgment to decide that Joseph Saikewicz, an adult who had never been competent, would have refused treatment had he been competent. Acknowledging that what the majority of reasonable people would choose might differ from the choice of a particular incompetent person, the court judiciously affirms, "The decision in many cases such as this should be that which would be made by the incompetent person, if that person were competent, but taking into account the present and future incompetency of the individual as one of the factors which would necessarily enter into the decision-making process of the competent person."[110]

This standard of substituted judgment could and should be used for once-competent patients, but only if reason exists to believe that the surrogate decision maker can make a judgment that the patient would have made.[111] In such cases, the surrogate should have a sufficiently deep familiarity with the patient that the particular judgment made reflects the patient's views and values. Merely knowing something in general about the patient's personal values is not sufficient. Accordingly, if the surrogate can reliably answer the question, "What would *the patient* want in this circumstance?" substituted judgment is an appropriate standard that approximates first-person consent. However, if the surrogate can only answer the question, "What do *you* want for the patient?" then a choice should be made on the basis of the patient's best interests rather than an autonomy standard. We cannot follow a substituted judgment standard for never-competent patients, because no basis exists for a judgment of their autonomous choice.

The Pure Autonomy Standard

A second standard eliminates the questionable idea of autonomy in the substituted judgment standard and replaces it with real autonomy. The pure autonomy standard applies exclusively to formerly autonomous, now-incompetent patients who, when autonomous, expressed a relevant treatment preference. The principle of respect for autonomy morally compels us to respect such clear preferences, even if the person can no longer express the preference for himself or herself. Whether or not a formal advance directive exists, this standard holds that caretakers should act on the patient's prior autonomous judgments, sometimes called "precedent autonomy."

Disputes arise, however, about the criteria of satisfactory evidence to support taking action under this standard. In the absence of explicit instructions, a surrogate decision maker might select from the patient's life history values that accord with the surrogate's own values, and then use only those values in reaching decisions. The surrogate might also base his or her findings on the patient's values that are only distantly relevant to the immediate decision (e.g., the patient's expressed dislike of hospitals). It is reasonable to ask what a surrogate decision maker can legitimately infer from a patient's prior conduct, especially from conditions such as fear and avoidance of doctors and earlier refusals to consent to physician recommendations.

Some evidence has been collected that surrogate decision makers for hospitalized older adults focus more on the patients' best interests than on the patients' prior preferences unless those preferences were explicitly formulated in advance directives.[112] Of course, even when the patient has provided an oral or written advance directive, surrogates need to determine whether it displays an autonomous preference that is directly pertinent to the decision at hand.[113]

The Best Interests Standard

Often a patient's relevant autonomous preferences cannot be determined. Under the best interests standard, a surrogate decision maker must then determine the highest probable net benefit among the available options, assigning different weights to interests the patient has in each option balanced against their inherent risks, burdens, or costs. The term *best* applies because of the surrogate's obligation to act beneficently by maximizing benefit through a comparative assessment that locates the highest probable net benefit. The best interests standard protects an incompetent person's welfare interests by requiring surrogates to assess the risks and probable benefits of various treatments and alternatives to treatment. It is therefore inescapably a quality-of-life criterion.

The best interests standard can justifiably override consents or refusals by minors or other incompetent patients, but, less obviously, it can also, in some circumstances, justifiably override advance directives appropriately prepared by formerly autonomous patients. This overriding can occur, for example, in a case in which a person by durable power of attorney has designated a surrogate to make medical decisions on his or her behalf. If the designated surrogate makes a decision that threatens the patient's best interests, the decision morally can and should be overridden by the medical team unless the patient while competent executed a clearly worded document that specifically supports the surrogate's decision.

Challenges to reliance on advance directives often stress the formerly autonomous person's failure to anticipate the circumstances that emerged. Examples are cases of apparently contented, nonsuffering, incompetent patients who can be expected to survive if treated against their advance directive but who otherwise would die. Discussions in the relevant literature at one time focused on the case of "Margo," a patient with Alzheimer's who, according to a medical student who visited her regularly, is "one of the happiest people I have ever known."[114] Some discussants ask us to imagine what should be done if Margo had a living will, executed just at the onset of her Alzheimer's, stating that she did not want life-sustaining treatment if she developed another life-threatening illness. In that circumstance caregivers would have to determine whether to honor her advance directive, and thereby to respect her precedent autonomy, by not using antibiotics to treat her pneumonia, or to act in accord with what may appear to be her current best interests in light of her overall happiness.

As persons slip into incompetence, their condition can be very different from, and sometimes better than, they had anticipated. If so, it seems unfair to the now happily situated incompetent person to be bound by a prior decision that may have been underinformed and shortsighted. In Margo's case, not using antibiotics would arguably harm what Ronald Dworkin calls, in discussing her case, her "experiential interests"—that is, her contentment with her current life. However, providing antibiotics would violate her living will, which expresses her considered values, her life story and commitments, and the like. Dworkin argues that Margo therefore should not be treated in these circumstances.[115] By contrast, the President's Council on Bioethics concluded that "Margo's apparent happiness would seem to make the argument for overriding the living will morally compelling in this particular case."[116]

Except in unusual cases, such as Margo's, we are obligated to respect the previously expressed autonomous wishes of the now-nonautonomous person because of the continuing force of the principle of respect for the autonomy of the person who made the decision. However, as we have seen, advance directives raise complex issues and occasionally can be justifiably overridden.

In this section we have argued that previously competent patients who autonomously expressed clear preferences in an oral or written advance directive should be treated under the pure autonomy standard, and we have suggested an economy of standards by viewing the first standard (substituted judgment) and the second standard (pure autonomy) as essentially identical. However, if the previously competent person left no reliable trace of his or her preferences—or if the individual was never competent—surrogate decision makers should adhere to the best interests standard.

CONCLUSION

The intimate connection between autonomy and decision making in health care and research, notably in circumstances of consent and refusal, unifies this chapter's several sections. We have justified the obligation to solicit decisions from patients and potential research subjects by appeal to the principle of respect for autonomy, but we have also acknowledged that the principle's precise demands can require thoughtful, and sometimes meticulous, interpretation and specification.

We have criticized various approaches to obtaining consents, but we are mindful that the history of informed consent and the place of autonomy in biomedical ethics are still under development. Current deficiencies in our systems and practices may become apparent in the near future just as we now recognize the past moral failures noted in this chapter. In examining standards for surrogate decision makers to use in regard to nonautonomous

patients, we have proposed an integrated set of standards of (1) respect for the patient's prior autonomous choices where reliably known and (2) the patient's best interests in the absence of reliable knowledge of the patient's prior autonomous choices. We have argued that (2) occasionally justifiably overrides (1) in circumstances of a conflict between the two.

We again stress in this conclusion that it is indefensible to construe respect for autonomy as a principle with priority over all other moral principles; it is one principle in our framework of prima facie principles suitable for biomedical ethics. The human moral community—indeed, morality itself—is rooted no less deeply in the three clusters of principles to be discussed in the next three chapters.

NOTES

1. 1. Those who enroll in research are generally referred to as *subjects*, but occasionally as *participants*. The choice of words can be morally significant. See the discussion of this distinction in National Bioethics Advisory Commission (NBAC), *Ethical and Policy Issues in Research Involving Human Participant*s, vol. I, *Report and Recommendations* (Bethesda, MD: NBAC, August 2001), pp. 32–33. See also our Chapter 6, endnote 1.
2. 2. The core idea of autonomy is treated by Joel Feinberg, *Harm to Self*, vol. 3 in *The Moral Limits of Criminal Law* (New York: Oxford University Press, 1986), chaps. 18–19; various essays in Franklin G. Miller and Alan Wertheimer, eds., *The Ethics of Consent: Theory and Practice* (New York: Oxford University Press, 2010); and several essays in James Stacey Taylor, ed., *Personal Autonomy: New Essays on Personal Autonomy and Its Role in Contemporary Moral Philosophy* (Cambridge: Cambridge University Press, 2005).
3. 3. For an argument that points to the importance of developing a broader theory of the nature of autonomy than we provide, see Rebecca Kukla, "Conscientious Autonomy: Displacing Decisions in Health Care," *Hastings Center Report* 35 (March–April 2005): 34–44; and Kukla, "Living with Pirates: Common Morality and Embodied Practice," *Cambridge Quarterly of Healthcare Ethics* 23 (2014): 75–85.
4. 4. Gerald Dworkin, *The Theory and Practice of Autonomy* (New York: Cambridge University Press, 1988), chaps. 1–4; Harry G. Frankfurt, "Freedom of the Will and the Concept of a Person," *Journal of Philosophy* 68 (1971): 5–20, as reprinted in *The Importance of What We Care About* (Cambridge: Cambridge University Press, 1988), pp. 11–25. Frankfurt may be primarily focused on a theory of freedom rather than a theory of autonomy; but see his uses of the language of "autonomy" in his *Necessity, Volition, and Love* (Cambridge: Cambridge University Press, 1999), chaps. 9, 11, especially pp. 95–110, 137.
5. 5. Dworkin, *The Theory and Practice of Autonomy*, p. 20.
6. 6. Agnieszka Jaworska, "Caring, Minimal Autonomy, and the Limits of Liberalism," in *Naturalized Bioethics: Toward Responsible Knowing and Practice*, ed. Hilde Lindemann, Marian Verkerk, and Margaret Urban Walker (New York: Cambridge University Press, 2009), pp. 80–105, esp. 82.
7. 7. For a "planning theory" and its relation to theories of autonomy, see Michael Bratman, "Planning Agency, Autonomous Agency," in *Personal Autonomy*, ed. Taylor, pp. 33–57.
8. 8. See the issues identified in Arthur Kuflik, "The Inalienability of Autonomy," *Philosophy & Public Affairs* 13 (1984): 271–98; Joseph Raz, "Authority and Justification," *Philosophy & Public Affairs* 14 (1985): 3–29; and Christopher McMahon, "Autonomy and Authority," *Philosophy & Public Affairs* 16 (1987): 303–28.
9. 9. See several essays in *Relational Autonomy: Feminist Perspectives on Autonomy, Agency, and the Social Self*, ed. Catriona Mackenzie and Natalie Stoljar (New York: Oxford University Press, 2000); Natalie Stoljar, "Feminist Perspectives on Autonomy," *Stanford Encyclopedia of Philosophy* (Fall 2015 Edition), ed. Edward N. Zalta, available at https://plato.stanford.edu/archives/fall2015/entries/feminism-autonomy/ (retrieved May 2, 2018); Marilyn Friedman, *Autonomy, Gender, and Politics* (New York: Oxford University Press, 2003); Friedman, "Autonomy and Social Relationships: Rethinking the Feminist Critique," in Diana T. Meyers, ed., *Feminists Rethink the Self* (Boulder, CO: Westview Press, 1997), pp. 40–61; Jennifer K. Walter and Lainie Friedman Ross, "Relational Autonomy: Moving beyond the Limits of Isolated Individualism," *Pediatrics* 133, Supplement 1 (2014): S16–S23; and Alasdair Maclean on

"relational consent" in his *Autonomy, Informed Consent and Medical Law: A Relational Challenge* (Cambridge: Cambridge University Press, 2009). See also the analysis of relational autonomy in James F. Childress, "Autonomy" [Addendum], *Bioethics* (formerly *Encyclopedia of Bioethics*), 4th ed., editor in chief, Bruce Jennings (Farmington Hills, MI: Gale, Cengage Learning—Macmillan Reference USA, 2014), vol. 1, pp. 307–9.

10. 10. See, further, Natalie Stoljar, "Informed Consent and Relational Conceptions of Autonomy," *Journal of Medicine and Philosophy* 36 (2011): 375–84; Carolyn Ells, "Shifting the Autonomy Debate to Theory as Ideology," *Journal of Medicine and Philosophy* 26 (2001): 417–30; Susan Sherwin, "A Relational Approach to Autonomy in Health-Care," in *The Politics of Women's Health: Exploring Agency and Autonomy*, The Feminist Health Care Ethics Research Network (Philadelphia: Temple University Press, 1998); and Anne Donchin, "Understanding Autonomy Relationally," *Journal of Medicine and Philosophy* 23, no. 4 (1998).
11. 11. See Barbara Herman, "Mutual Aid and Respect for Persons," *Ethics* 94 (July 1984): 577–602, esp. 600–602; and Onora O'Neill, "Universal Laws and Ends-in-Themselves," *Monist* 72 (1989): 341–61.
12. 12. This misunderstanding of our views is found in M. Therese Lysaught, "Respect: or, How Respect for Persons Became Respect for Autonomy," *Journal of Medicine and Philosophy* 29 (2004): 665–80, esp. 676.
13. 13. Carl E. Schneider, *The Practice of Autonomy: Patients, Doctors, and Medical Decisions* (New York: Oxford University Press, 1998), esp. p. xi. For various views supportive of a limited role for the principle of respect for autonomy, see Paul Root Wolpe, "The Triumph of Autonomy in American Bioethics: A Sociological View," in *Bioethics and Society: Constructing the Ethical Enterprise*, ed. Raymond DeVries and Janardan Subedi (Upper Saddle River, NJ: Prentice Hall, 1998), pp. 38–59; Sarah Conly, *Against Autonomy: Justifying Coercive Paternalism* (Cambridge: Cambridge University Press, 2013); Jukka Varelius, "The Value of Autonomy in Medical Ethics," *Medicine, Health Care, and Philosophy* 9 (2006): 377–88; Daniel Callahan, "Autonomy: A Moral Good, Not a Moral Obsession," *Hastings Center Report* 14 (October 1984): 40–42. Contrast James F. Childress, "The Place of Autonomy in Bioethics," *Hastings Center Report* 20 (January–February 1990): 12–16; and Thomas May, "The Concept of Autonomy in Bioethics: An Unwarranted Fall from Grace," in *Personal Autonomy*, ed. Taylor, pp. 299–309.
14. 14. Leslie J. Blackhall, Sheila T. Murphy, Gelya Frank, et al., "Ethnicity and Attitudes toward Patient Autonomy," *JAMA: Journal of the American Medical Association* 274 (September 13, 1995): 820–25.
15. 15. Joseph A. Carrese and Lorna A. Rhodes, "Western Bioethics on the Navajo Reservation: Benefit or Harm?" *JAMA: Journal of the American Medical Association* 274 (September 13, 1995): 826–29.
16. 16. We make these points to forestall misunderstanding. Some critics of theories that connect respect for autonomy to informed consent mistakenly presume that defenders of these views, including us, view consent as necessary and sufficient. See, for example, Neil C. Manson and Onora O'Neill, *Rethinking Informed Consent in Bioethics* (Cambridge: Cambridge University Press, 2007), pp. 19, 185ff.
17. 17. For further discussion of the relation between autonomy and consent, see Tom L. Beauchamp, "Autonomy and Consent," in *The Ethics of Consent*, ed. Miller and Wertheimer, chap. 3.
18. 18. See Avram Goldstein, "Practice vs. Privacy on Pelvic Exams," *Washington Post*, May 10, 2003, p. A1, available at https://www.washingtonpost.com/archive/politics/2003/05/10/practice-vs-privacy-on-pelvic-exams/4e9185c4-4b4c-4d6a-a132-b21b8471da58/?utm_term=.ee1d008b73ce (accessed May 8, 2018).
19. 19. For studies of views of women in Canada and Ireland, see S. Wainberg, H. Wrigley, J. Fair, and S. Ross, "Teaching Pelvic Examinations under Anaesthesia: What Do Women Think?" *Journal of Obstetrics and Gynaecology Canada*, *Journal d'Obstétrique et Gynécologie du Canada* 32, no. 1 (2010): 49–53; and F. Martyn and R. O'Connor, "Written Consent for Intimate Examinations Undertaken by Medical Students in the Operating Theatre—Time for National Guidelines?" *Irish Medical Journal* 102, no. 10 (2009): 336–37. See also the discussion of the evidence about women's views in Phoebe Friesen, "Educational Pelvic Exams on Anesthetized Women: Why Consent Matters," *Bioethics* 32 (2018): 298–307.
20. 20. Britt-Ingjerd Nesheim, "Commentary: Respecting the Patient's Integrity Is the Key," *BMJ: British Medical Journal* 326 (January 11, 2003): 100. For a thorough examination of the ethical issues and an argument that the practice of unconsented pelvic examinations in medical education is "immoral and indefensible," see Friesen, "Educational Pelvic Exams on Anesthetized Women: Why Consent Matters."
21. 21. See Shawn S. Barnes, "Practicing Pelvic Examinations by Medical Students on Women under Anesthesia: Why Not Ask First?" *Obstetrics and Gynecology* 120, no. 4 (2012): 941–43; and Arthur L.

Caplan, "Pelvic Exams Done on Anesthetized Women without Consent: Still Happening," *Medscape*, May 2, 2018, available at https://www.medscape.com/viewarticle/894693 (accessed October 7, 2018).

22. 22. Peter A. Ubel, Christopher Jepson, and Ari Silver-Isenstadt, "Don't Ask, Don't Tell: A Change in Medical Student Attitudes after Obstetrics/Gynecology Clerkships toward Seeking Consent for Pelvic Examinations on an Anesthetized Patient," *American Journal of Obstetrics and Gynecology* 188 (February 2003): 575–79.
23. 23. Bernard M. Branson, H. Hunter Handsfield, Margaret A. Lampe, et al., "Revised Recommendations for HIV Testing of Adults, Adolescents, and Pregnant Women in Health-Care Settings," *Morbidity and Mortality Weekly Report, Recommendations and Report* 55 (RR-14) (September 22, 2006): 1–17. These recommendations expect specific, explicit informed consent in nonclinical settings.
24. 24. See Ronald Bayer and Amy L. Fairchild, "Changing the Paradigm for HIV Testing—The End of Exceptionalism," *New England Journal of Medicine* 355 (August 17, 2006): 647–49; Lawrence O. Gostin, "HIV Screening in Health Care Settings: Public Health and Civil Liberties in Conflict?" *JAMA: Journal of the American Medical Association* 296 (October 25, 2006): 2023–25; and Thomas R. Frieden et al., "Applying Public Health Principles to the HIV Epidemic," *New England Journal of Medicine* 353 (December 1, 2005): 2397–402. For a cost-effectiveness analysis, see Gillian D. Sanders et al., "Cost-Effectiveness of Screening for HIV in the Era of Highly Active Antiretroviral Therapy," *New England Journal of Medicine* 352 (February 10, 2005): 570–85.
25. 25. See HIVgov, *U.S. Statistics*, available at https://www.hiv.gov/hiv-basics/overview/data-and-trends/statistics (accessed October 12, 2018).
26. 26. See Centers for Disease Control and Prevention, HIV/AIDS, *HIV Treatment as Prevention*, available at https://www.cdc.gov/hiv/risk/art/index.html (accessed October 11, 2018); and Myron S. Cohen and Cynthia L. Gay, "Treatment to Prevent Transmission of HIV-1," *Clinical Infectious Diseases 50 (2010): S85–S95. See also Carl W. Dieffenbach and Anthony S. Fauci, "Thirty Years of HIV and AIDS: Future Challenges and Opportunities," Annals of Internal Medicine* 154, no. 11 (June 2011): 766–72.
27. 27. Centers for Disease Control and Prevention, HIV/AIDS, *HIV Treatment as Prevention*.
28. 28. Quoted in Bayer and Fairchild, "Changing the Paradigm for HIV Testing," p. 649.
29. 29. For the evolution of informed consent in HIV testing in the United States, with attention to several factors that led to the end of written informed consent, see Ronald Bayer, Morgan Philbin, and Robert H. Remien, "The End of Written Informed Consent for HIV Testing: Not with a Bang but a Whimper," *American Journal of Public Health* 107, no. 8 (August 2017): 1259–65. Nebraska, the last state to change its law, did so in 2018 after this article appeared. See Nebraska Legislature, *Legislative Bill 285* (Approved by the governor February 28, 2018), available at https://nebraskalegislature.gov/FloorDocs/105/PDF/Slip/LB285.pdf (accessed October 7, 2018).
30. 30. For a comprehensive discussion of the issues raised by "opt-out" policies to increase the supply of transplantable organs, see J. Bradley Segal and Robert D. Truog, "Options for Increasing the Supply of Transplantable Organs," *Harvard Health Policy Review*, December 2, 2017, available at http://www.hhpronline.org/articles/2017/12/2/options-for-increasing-the-supply-of-transplantable-organs-2 (accessed May 2, 2018); and Institute of Medicine (now Academy of Medicine), Committee on Increasing Rates of Organ Donation, *Organ Donation: Opportunities for Action*, ed. James F. Childress and Catharyn Liverman (Washington, DC: National Academies Press, 2006), chap. 7. See also Richard H. Thaler and Cass R. Sunstein, *Nudge: Improving Decisions about Health, Wealth, and Happiness* (New Haven, CT: Yale University Press, 2008), chap. 11, "How to Increase Organ Donations."
31. 31. This case was developed by Dr. Gail Povar.
32. 32. See Thomas Grisso and Paul S. Appelbaum, *Assessing Competence to Consent to Treatment: A Guide for Physicians and Other Health Professionals* (New York: Oxford University Press, 1998), p. 11.
33. 33. The analysis in this section has profited from discussions with Ruth R. Faden, Nancy M. P. King, and Dan Brock.
34. 34. See the examination of the core meaning in Charles M. Culver and Bernard Gert, *Philosophy in Medicine* (New York: Oxford University Press, 1982), pp. 123–26.
35. 35. *Pratt v. Davis*, 118 Ill. App. 161 (1905), aff'd, 224 Ill. 300, 79 N.E. 562 (1906).
36. 36. See Daniel Wikler, "Paternalism and the Mildly Retarded," *Philosophy & Public Affairs* 8 (1979): 377–92; and Kenneth F. Schaffner, "Competency: A Triaxial Concept," in *Competency*, ed. M. A. G. Cutter and E. E. Shelp (Dordrecht, Netherlands: Kluwer Academic, 1991), pp. 253–81.

37. 37. This case was prepared by Dr. P. Browning Hoffman for presentation in the series of "Medicine and Society" conferences at the University of Virginia.
38. 38. Laura L. Sessums, Hanna Zembrzuska, and Jeffrey L. Jackson, "Does This Patient Have Medical Decision-Making Capacity?" *JAMA: Journal of the American Medical Association* 306 (July 27, 2011): 420–27. See also J. B. Jourdan and L. Glickman, "Reasons for Requests for Evaluation of Competency in a Municipal General Hospital," *Psychomatics* 32 (1991): 413–16.
39. 39. This schema is indebted to Paul S. Appelbaum and Thomas Grisso, "Assessing Patients' Capacities to Consent to Treatment," *New England Journal of Medicine* 319 (December 22, 1988): 1635–38; Appelbaum and Grisso, "The MacArthur Treatment Competence Study I. Mental Illness and Competence to Consent to Treatment," *Law and Human Behavior* 19 (1995): 105–26; and Jessica W. Berg, Paul S. Appelbaum, Charles W. Lidz, and Lisa S. Parker, *Informed Consent: Legal Theory and Clinical Practice*, 2nd ed. (New York: Oxford University Press, 2001).
40. 40. For a comprehensive treatment, see Ian McDowell, *Measuring Health: A Guide to Rating Scales and Questionnaires*, 3rd ed. (Oxford: Oxford University Press, 2006).
41. 41. For additional ways in which values are incorporated, see Loretta M. Kopelman, "On the Evaluative Nature of Competency and Capacity Judgments," *International Journal of Law and Psychiatry* 13 (1990): 309–29. For conceptual and epistemic problems in available tests, see E. Haavi Morreim, "Competence: At the Intersection of Law, Medicine, and Philosophy," in *Competency*, ed. Cutter and Shelp, pp. 93–125, esp. pp. 105–8.
42. 42. It is beyond the scope of our discussion to analyze and evaluate the numerous tests and instruments that have been developed to assess decisional capacity for clinical treatment and research. The following three books offer guidance to "best practices" of assessing competence: Grisso and Appelbaum, *Assessing Competence to Consent to Treatment: A Guide for Physicians and Other Health* Professionals; Scott Y. H. Kim, *Evaluation of Capacity to Consent to Treatment and Research*, Best Practices in Forensic Mental Health Assessment (New York: Oxford University Press, 2010); and Deborah Bowman, John Spicer, and Rehana Iqbal, *Informed Consent: A Primer for Clinical Practice* (Cambridge: Cambridge University Press, 2012), chapter 2, "On Capacity: Can the Patient Decide?"
43. 43. Grisso and Appelbaum, *Assessing Competence to Consent to Treatment*, p. 139.
44. 44. Allen Buchanan and Dan Brock, *Deciding for Others* (Cambridge: Cambridge University Press, 1989), pp. 51–70; Willard Gaylin, "The Competence of Children: No Longer All or None," *Hastings Center Report* 12 (1982): 33–38, esp. 35; and Eric Kodish, "Children's Competence for Assent and Consent: A Review of Empirical Findings," *Ethics & Behavior* 14 (2004): 255–95.
45. 45. Buchanan and Brock, *Deciding for Others*, pp. 52–55. For elaboration and defense, see Brock, "Decisionmaking Competence and Risk," *Bioethics* 5 (1991): 105–12.
46. 46. NBAC, *Report and Recommendations of the National Bioethics Advisory Commission, Research Involving Persons with Mental Disorders That May Affect Decision Making Capacity*, vol. 1 (Rockville, MD: National Bioethics Advisory Commission, December 1998), p. 58.
47. 47. For concise accounts of how informed consent grew and developed in law, regulation, and policy, principally in the United States, see Alexander M. Capron, "Legal and Regulatory Standards of Informed Consent in Research," in *The Oxford Textbook of Clinical Research Ethics*, ed. Ezekiel Emanuel, Christine Grady, Robert Crouch, et al. (New York: Oxford University Press, 2008), pp. 613–32; Presidential Commission for the Study of Bioethical Issues, "Informed Consent Background" (as updated September 30, 2016), available at https://bioethicsarchive.georgetown.edu/pcsbi/sites/default/files/1%20Informed%20Consent%20Background%209.30.16.pdf (accessed May 6, 2018); and Faden and Beauchamp, *A History and Theory of Informed Consent*, chaps. 2 and 4.
48. 48. See Neal W. Dickert, Nir Eyal, Sara F. Goldkind, et al., "Reframing Consent for Clinical Research: A Function-Based Approach," *American Journal of Bioethics* 17 (2017): 3–11. See the reply to these authors by Tom L. Beauchamp, "The Idea of a 'Standard View' of Informed Consent," *American Journal of Bioethics* 17 (2017): 1–2 (editorial). For analysis of the justification of informed consent in research, see Dan W. Brock, "Philosophical Justifications of Informed Consent in Research," in *The Oxford Textbook of Clinical Research Ethics*, ed. Emanuel, Grady, Crouch, et al., pp. 606–12. Brock is a coauthor of "Reframing Consent for Clinical Research: A Function-Based Approach," and his work implicitly shows

the compatibility of a function-based approach with one grounded in normative philosophical justifications.

49. 49. Onora O'Neill, *Autonomy and Trust in Bioethics* (Cambridge: Cambridge University Press, 2002); O'Neill, "Autonomy: The Emperor's New Clothes," *Proceedings of the Aristotelian Society*, supp. vol. 77 (2003): 1–21; O'Neill, "Some Limits of Informed Consent," *Journal of Medical Ethics* 29 (2003): 4–7; and Manson and O'Neill, *Rethinking Informed Consent in Bioethics*.
50. 50. O'Neill, "Some Limits of Informed Consent," p. 5.
51. 51. See Jay Katz, *The Silent World of Doctor and Patient* (New York: Free Press, 1984), pp. 86–87 (Reprint ed. Baltimore, MD: Johns Hopkins University Press, 2002); and President's Commission for the Study of Ethical Problems in Medicine and Biomedical and Behavioral Research, *Making Health Care Decisions*, vol. 1 (Washington, DC: US Government Printing Office, 1982), p. 15.
52. 52. See James F. Childress, "Needed: A More Rigorous Analysis of Models of Decision Making and a Richer Account of Respect for Autonomy," *American Journal of Bioethics* 17, no. 11 (2017): 52–54, in response to Peter A. Ubel, Karen A. Scherr, and Angela Fagerlin, "Empowerment Failure: How Shortcomings in Physician Communication Unwittingly Undermine Patient Autonomy," *American Journal of Bioethics* 17, no. 11 (2017): 31–39, which seeks to combine a model of shared decision making with patient empowerment. See, in turn, Ubel, Scherr, and Fagerlin, "Autonomy: What's Shared Decision Making Have to Do with It?" *American Journal of Bioethics* 18, no. 2 (February 2018): W11–W12, which concedes the problems with the term "shared decision making," but stresses that it refers to the "process" of decision making and could be called "assisted decision making" and argues, less convincingly, that challenging the legitimacy of the increasingly accepted term at this point could actually damage patient autonomy.
53. 53. For extensions of this thesis, see Simon Whitney, Amy McGuire, and Laurence McCullough, "A Typology of Shared Decision Making, Informed Consent, and Simple Consent," *Annals of Internal Medicine* 140 (2004): 54–59.
54. 54. The analysis in this subsection is based in part, but substantially, on Faden and Beauchamp, *A History and Theory of Informed Consent*, chap. 8.
55. 55. *Mohr v. Williams*, 95 Minn. 261, 265; 104 N.W. 12, at 15 (1905).
56. 56. Franklin G. Miller and Alan Wertheimer, "The Fair Transaction Model of Informed Consent: An Alternative to Autonomous Authorization," *Kennedy Institute of Ethics Journal* 21 (2011): 201–18. On pp. 210–12 these authors recognize the importance of our second sense of "informed consent" and the qualifications it allows, but they do not confront our views about the critical importance of maintaining the first sense as the primary model of an informed consent. See further their "Preface to a Theory of Consent Transactions: Beyond Valid Consent," in *The Ethics of Consent*, ed. Miller and Wertheimer, pp. 79–105. For an expanded and revised version of the last essay, see Alan Wertheimer, *Rethinking the Ethics of Clinical Research: Widening the Lens* (New York: Oxford University Press, 2011), chap. 3.
57. 57. See, for example, National Commission for the Protection of Human Subjects of Biomedical and Behavioral Research, *The Belmont Report* (Washington, DC: DHEW Publication OS 78–0012, 1978), p. 10; Alexander M. Capron, "Legal and Regulatory Standards of Informed Consent in Research," pp. 623–32; Dan W. Brock, "Philosophical Justifications of Informed Consent in Research," pp. 607–11; Alan Meisel and Loren Roth, "What We Do and Do Not Know about Informed Consent," *JAMA: Journal of the American Medical Association* 246 (1981): 2473–77; and President's Commission, *Making Health Care Decisions*, vol. 2, pp. 317–410, esp. p. 318, and vol. 1, chap. 1, esp. pp. 38–39.
58. 58. A classic case is United States Supreme Court, *Planned Parenthood of Central Missouri v. Danforth*, 428 U.S. 52 at 67 n.8 (1976).
59. 59. See Capron, "Legal and Regulatory Standards of Informed Consent in Research," pp. 623–28.
60. 60. *Moore v. Regents of the University of California*, 793 P.2d 479 (Cal. 1990) at 483.
61. 61. See, for example, Clarence H. Braddock et al., "How Doctors and Patients Discuss Routine Clinical Decisions: Informed Decision Making in the Outpatient Setting," *Journal of General Internal Medicine* 12 (1997): 339–45; and John Briguglio et al., "Development of a Model Angiography Informed Consent Form Based on a Multiinstitutional Survey of Current Forms," *Journal of Vascular and Interventional Radiology* 6 (1995): 971–78.
62. 62. The subjective standard requires a physician to disclose the information a particular patient needs to know to the extent it is reasonable to expect the physician to be able to determine that patient's

informational needs. The Oklahoma Supreme Court supported this standard in *Scott v. Bradford*, 606 P.2d 554 (Okla. 1979) at 559 and *Masquat v. Maguire*, 638 P.2d 1105, Okla. 1981. For a defense of the subjective standard as the normative ethical ideal, see Vilius Dranseika, Jan Piasecki, and Marcin Waligora, "Relevant Information and Informed Consent in Research: In Defense of the Subjective Standard of Disclosure," *Science and Engineering Ethics* 23, no. 1 (2017): 215–25.

63. 63. Robert D. Truog, Walter Robinson, Adrienne Randolph, and Alan Morris, "Is Informed Consent Always Necessary for Randomized, Controlled Trials?" Sounding Board, *New England Journal of Medicine* 340 (March 11, 1999): 804–7; and Ruth R. Faden, Tom L. Beauchamp, and Nancy E. Kass, "Informed Consent, Comparative Effectiveness, and Learning Health Care," *New England Journal of Medicine* 370 (Feb. 20, 2014): 766-68.
64. 64. The literature on the ethical controversy about informed consent in the SUPPORT study is extensive. For an introduction to the issues, see *American Journal of Bioethics* 13, no. 12 (2013): 1526–61, particularly David Magnus, "The SUPPORT Controversy and the Debate over Research within the Standard of Care"; David Wendler, "What Should Be Disclosed to Research Participants?"; Ruth Macklin and Lois Shepherd, "Informed Consent and Standard of Care: What Must Be Disclosed"; and Benjamin S. Wilfond, "Quality Improvement Ethics: Lessons from the SUPPORT Study," along with several responses.
65. 65. *Canterbury v. Spence*, 464 F.2d 772 (1977), at 785–89; and see Nathan A. Bostick, Robert Sade, John W. McMahon, and Regina Benjamin, "Report of the American Medical Association Council on Ethical and Judicial Affairs: Withholding Information from Patients: Rethinking the Propriety of 'Therapeutic Privilege,'" *Journal of Clinical Ethics* 17 (Winter 2006): 302–6, pdf available at https://www.researchgate.net/publication/6475405_Report_of_the_American_Medical_Association_Council_on_Ethical_and_Judicial_Affairs_withholding_information_from_patients_rethinking_the_propriety_of_therapeutic_privilege (accessed May 7, 2018). For studies of levels of anxiety and stress produced by informed consent disclosures, see Jeffrey Goldberger et al., "Effect of Informed Consent on Anxiety in Patients Undergoing Diagnostic Electrophysiology Studies," *American Heart Journal* 134 (1997): 119–26; and Kenneth D. Hopper et al., "The Effect of Informed Consent on the Level of Anxiety in Patients Given IV Contrast Material," *American Journal of Roentgenology* 162 (1994): 531–35.
66. 66. *Thornburgh v. American College of Obstetricians*, 476 U.S. 747 (1986) (White, J., dissenting).
67. 67. For a report congenial to our conclusion, see Bostick, Sade, McMahon, and Benjamin, "Report of the American Medical Association Council on Ethical and Judicial Affairs: Withholding Information from Patients: Rethinking the Propriety of 'Therapeutic Privilege,'" pp. 302–6. The term *therapeutic privilege* does not appear in the current AMA Code. See *Code of Medical Ethics of the American Medical Association*, 2016–2017 Edition (Chicago: AMA, 2017), 2.1.3, "Withholding Information from Patients." This code stresses dispensing information in accord with patients' preferences and hence their autonomous choices.
68. 68. Asbjørn Hróbjartsson and Peter C Gøtzsche, "Placebo Interventions for All Clinical Conditions (Review)," The Cochrane Collaboration (Chichester, UK: John Wiley, 2010), available at https://nordic.cochrane.org/sites/nordic.cochrane.org/files/public/uploads/ResearchHighlights/Placebo%20interventions%20for%20all%20clinical%20conditions%20(Cochrane%20review).pdf (accessed October 11, 2018).
69. 69. Howard Brody, *Placebos and the Philosophy of Medicine: Clinical, Conceptual, and Ethical Issues* (Chicago: University of Chicago Press, 1980), pp. 10–11.
70. 70. Ted J. Kaptchuk, Elizabeth Friedlander, John M. Kelley, et al., "Placebos without Deception: A Randomized Controlled Trial in Irritable Bowel Syndrome," *PLOS One* 5 (2010), available at http://www.plosone.org/article/info:doi/10.1371/journal.pone.0015591 (accessed October 11, 2018).
71. 71. Michael E. Wechsler, John M. Kelley, Ingrid O. E. Boyd, et al., "Active Albuterol or Placebo, Sham Acupuncture, or No Intervention in Asthma," *New England Journal of Medicine* 365 (July 14, 2011): 119–26.
72. 72. Jon C. Tilburt, Ezekiel J. Emanuel, Ted J. Kaptchuk, et al., "Prescribing 'Placebo Treatments': Results of National Survey of US Internists and Rheumatologists," *BMJ* 337 (2008): a1938. Similar results have been reported in studies in other countries. See, for example, Corey S. Harris, Natasha K. J. Campbell, and Amir Raz, "Placebo Trends across the Border: US versus Canada," *PLOS One* 10, no. 11 (2015):

e0142804; and J. Howick, F. L. Bishop, C. Heneghan, et al., "Placebo Use in the United Kingdom: Results from a National Survey of Primary Care Practitioners," *PLOS One* 8, no. 3 (2013): e58247.

73. 73. Sara Chandros Hull, Luana Colloca, Andrew Avins, et al., "Patients' Attitudes about the Use of Placebo Treatments: Telephone Survey," *BMJ* 347 (2013); f3757. Most also favored transparency and honesty. The place and ethics of placebos in medicine have also received considerable attention in magazines for the public. See Michael Specter, "The Power of Nothing: Could Studying the Placebo Effect Change the Way We Think about Medicine?" *New Yorker*, December 12, 2011; and Elaine Schattner, "The Placebo Debate: Is It Unethical to Prescribe Them to Patients?" *Atlantic*, December 19, 2011.
74. 74. On the merit of these arguments, see Anne Barnhill, "What It Takes to Defend Deceptive Placebo Use," *Kennedy Institute of Ethics Journal* 21 (2011): 219–50. See also Sissela Bok, "Ethical Issues in Use of Placebo in Medical Practice and Clinical Trials," in *The Science of the Placebo: Toward an Interdisciplinary Research Agenda*, ed. Harry A. Guess, Arthur Kleinman, John W. Kusek, and Linda W. Engel (London: BMJ Books, 2002), pp. 53–74.
75. 75. For a similar proposal, see Armand Lione, "Ethics of Placebo Use in Clinical Care" (Correspondence), *Lancet* 362 (September 20, 2003): 999. For cases involving the different appeals to "consent," along with analysis and assessment, see P. Lichtenberg, U. Heresco-Levy, and U. Nitzan, "The Ethics of the Placebo in Clinical Practice," *Journal of Medical Ethics* 30 (2004): 551–54; and "Case Vignette: Placebos and Informed Consent," *Ethics and Behavior* 8 (1998): 89–98, with commentaries by Jeffrey Blustein, Walter Robinson, Gregory S. Loeben, and Benjamin S. Wilfond.
76. 76. *Code of Medical Ethics of the American Medical Association*, 2016–2017 Edition, 2.1.4, "Use of Placebo in Clinical Practice." For a criticism of an earlier, but somewhat similar, version of this policy, see both Bennett Foddy, "A Duty to Deceive: Placebos in Clinical Practice," *American Journal of Bioethics* 9, no. 12 (2009): 4–12 (and his response to commentaries in the same issue, W1–2); and Adam Kolber, "A Limited Defense of Clinical Placebo Deception," *Yale Law & Policy Review* 26 (2007): 75–134. For a defense of the earlier version, see Kavita R. Shah and Susan Door Goold, "The Primacy of Autonomy, Honesty, and Disclosure—Council on Ethical and Judicial Affairs' Placebo Opinions," *American Journal of Bioethics* 9, no. 12 (2009): 15–17. For an analysis of the science and ethics of placebo treatment, see Franklin G. Miller and Luana Colloca, "The Legitimacy of Placebo Treatments in Clinical Practice: Evidence and Ethics," *American Journal of Bioethics* 9, no. 12 (2009): 39–47; and Damien G. Finnis, Ted J. Kaptchuk, Franklin G. Miller, and Fabrizio Benedetti, "Biological, Clinical, and Ethical Advances of Placebo Effects," *Lancet* 375, no. 9715 (February 20, 2010): 696–95. See also N. Biller-Andorno, "The Use of the Placebo Effect in Clinical Medicine—Ethical Blunder or Ethical Imperative?" *Science and Engineering Ethics* 10 (2004): 43–50.
77. 77. Kaptchuk, Friedlander, Kelley, et al., "Placebos without Deception"; Brody, *Placebos and the Philosophy of Medicine*, pp. 110, 113, et passim; and Brody, "The Placebo Response: Recent Research and Implications for Family Medicine," *Journal of Family Practice* 49 (July 2000): 649–54. For a broad defense of placebos, see Howard Spiro, *Doctors, Patients, and Placebos* (New Haven, CT: Yale University Press, 1986).
78. 78. See Fabrizio Benedetti, "Mechanisms of Placebo and Placebo-Related Effects across Diseases and Treatments," *Annual Review of Pharmacology and Toxicology* 48 (2008): 33–60, and more fully developed in his *Placebo Effects: Understanding the Mechanisms in Health and Disease* (New York: Oxford University Press, 2009). Benedetti focuses on the "psychosocial-induced biochemical changes in a person's brain and body."
79. 79. See Yael Schenker, Alicia Fernandez, and Bernard Lo, "Placebo Prescriptions Are Missed Opportunities for Doctor-Patient Communication," *American Journal of Bioethics* 9 (2009): 48–50; and Howard Brody, "Medicine's Continuing Quest for an Excuse to Avoid Relationships with Patients," *American Journal of Bioethics* 9 (2009): 13–15.
80. 80. Sally E. McNagy and Ruth M. Parker, "High Prevalence of Recent Cocaine Use and the Unreliability of Patient Self-Report in an Inner-City Walk-in Clinic," *JAMA: Journal of the American Medical Association* 267 (February 26, 1992): 1106–8.
81. 81. Sissela Bok, "Informed Consent in Tests of Patient Reliability," *JAMA: Journal of the American Medical Association* 267 (February 26, 1992): 1118–19.

82. 82. Barbara A. Bernhardt et al., "Educating Patients about Cystic Fibrosis Carrier Screening in a Primary Care Setting," *Archives of Family Medicine* 5 (1996): 336–40; Leanne Stunkel, Meredith Benson, Louise McLellan, et al., "Comprehension and Informed Consent: Assessing the Effect of a Short Consent Form," *IRB* 32 (2010): 1–9; and James H. Flory, David Wendler, and Ezekiel J. Emanuel, "Empirical Issues in Informed Consent for Research," in *The Oxford Textbook of Clinical Research Ethics*, ed. Emanuel, Grady, Crouch, et al., pp. 645–60.
83. 83. Steven Joffe, E. Francis Cook, Paul D. Cleary, et al., "Quality of Informed Consent in Cancer Clinical Trials: A Cross-Sectional Survey," *Lancet* 358 (November 24, 2001): 1772–77. See further Joffe, Cook, Cleary, et al., "Quality of Informed Consent: A New Measure of Understanding among Research Subjects," *JNCI: Journal of the National Cancer Institute* 93 (January 17, 2001): 139–47; and Michael Jefford and Rosemary Moore, "Improvement of Informed Consent and the Quality of Consent Documents," *Lancet Oncology* 9 (2008): 485–93.
84. 84. *Bang v. Charles T. Miller Hospital*, 88 N.W. 2d 186, 251 Minn. 427, 1958 Minn.
85. 85. See further Gopal Sreenivasan, "Does Informed Consent to Research Require Comprehension?" *Lancet* 362 (December 13, 2003): 2016–18.
86. 86. C. K. Dougherty et al., "Perceptions of Cancer Patients and Their Physicians Involved in Phase I Clinical Trials," *Journal of Clinical Oncology* 13 (1995): 1062–72; and Paul R. Benson et al., "Information Disclosure, Subject Understanding, and Informed Consent in Psychiatric Research," *Law and Human Behavior* 12 (1988): 455–75.
87. 87. See further Edmund G. Howe, "Approaches (and Possible Contraindications) to Enhancing Patients' Autonomy," *Journal of Clinical Ethics* 5 (1994): 179–88.
88. 88. See Michael B. Rothberg, Senthil K. Sivalingam, Javed Ashraf, et al., "Patients' and Cardiologists' Perceptions of the Benefits of Percutaneous Coronary Intervention for Stable Coronary Disease," *Annals of Internal Medicine* 153 (2010): 307–13. See also the commentary by Alicia Fernandez, "Improving the Quality of Informed Consent: It Is Not All about the Risks," *Annals of Internal Medicine* 153 (2010): 342–43.
89. 89. This label was apparently coined by Paul S. Appelbaum, Loren Roth, and Charles W. Lidz in "The Therapeutic Misconception: Informed Consent in Psychiatric Research," *International Journal of Law and Psychiatry* 5 (1982): 319–29. See further Appelbaum, Lidz, and Thomas Grisso, "Therapeutic Misconception in Clinical Research: Frequency and Risk Factors," *IRB: Ethics and Human Research* 26 (2004): 1–8; Walter Glannon, "Phase I Oncology Trials: Why the Therapeutic Misconception Will Not Go Away," *Journal of Medical Ethics* 32 (2006): 252–55; Appelbaum and Lidz, "The Therapeutic Misconception," in *The Oxford Textbook of Clinical Research Ethics*, ed. Emanuel, Grady, Crouch, et al.; Rebecca Dresser, "The Ubiquity and Utility of the Therapeutic Misconception," *Social Philosophy and Policy* 19 (2002): 271–94; and Franklin G. Miller, "Consent to Clinical Research," in *The Ethics of Consent: Theory and Practice*, ed. Miller and Wertheimer, chap. 15. See also Inmaculada de Melo-Martín and Anita Ho, "Beyond Informed Consent: The Therapeutic Misconception and Trust," *Journal of Medical Ethics* 34 (2008): 202–5.
90. 90. A broader problem and one more difficult to address is that the frame of discourse in interactions between researchers and potential subjects may incorporate the therapeutic misconception. See Philip J. Candilis and Charles W. Lidz, "Advances in Informed Consent Research," in *The Ethics of Consent*, ed., Miller and Wertheimer, p. 334; David E. Ness, Scott Kiesling, and Charles W. Lidz, "Why Does Informed Consent Fail? A Discourse Analytic Approach," *Journal of the American Academy of Psychiatry and the Law* 37 (2009): 349–62.
91. 91. Sam Horng and Christine Grady, "Misunderstanding in Clinical Research: Distinguishing Therapeutic Misconception, Therapeutic Misestimation, and Therapeutic Optimism," *IRB: Ethics and Human Research* 25 (January–February 2003): 11–16; and see also Horng, Ezekiel Emanuel, Benjamin Wilfond, et al., "Descriptions of Benefits and Risks in Consent Forms for Phase 1 Oncology Trials," *New England Journal of Medicine* 347 (2002): 2134–40.
92. 92. The pioneering work was done by Amos Tversky and Daniel Kahneman. See "Choices, Values and Frames," *American Psychologist* 39 (1984): 341–50; and "The Framing of Decisions and the Psychology of Choice," *Science* 211 (1981): 453–58. See also Daniel Kahneman and Amos Tversky, eds., *Choices, Values, and Frames* (Cambridge: Cambridge University Press, 2000). On informed consent specifically, see Dennis J. Mazur and Jon F. Merz, "How Age, Outcome Severity, and Scale Influence General

Medicine Clinic Patients' Interpretations of Verbal Probability Terms," *Journal of General Internal Medicine* 9 (1994): 268–71.

93. 93. S. E. Eraker and H. C. Sox, "Assessment of Patients' Preferences for Therapeutic Outcomes," *Medical Decision Making* 1 (1981): 29–39; Barbara McNeil et al., "On the Elicitation of Preferences for Alternative Therapies," *New England Journal of Medicine* 306 (May 27, 1982): 1259–62.
94. 94. See A. M. O'Connor, C. L. Bennett, D. Stacey, et al., "Decision Aids for People Facing Health Treatment or Screening Decisions," *Cochrane Database of Systematic Reviews*, no. 3 (2009), Art. No. CD001431; Philip J. Candilis and Charles W. Lidz, "Advances in Informed Consent Research," chap. 13; and Barton W. Palmer, Nicole M. Lanouette, and Dilip V. Jeste, "Effectiveness of Multimedia Aids to Enhance Comprehension of Research Consent Information: A Systematic Review," *IRB: Ethics & Human Research* 34 (2012), available at https://www.thehastingscenter.org/wp-content/uploads/nov-dec12irb-palmer-tables.pdf (accessed May 8, 2018).
95. 95. Ruth Faden and Alan Faden, "False Belief and the Refusal of Medical Treatment," *Journal of Medical Ethics* 3 (1977): 133–36.
96. 96. Neil C. Manson and Onora O'Neill interpret all consent as a waiver of rights. This interpretation is in some respects correct, but it is more illuminating in most cases to describe informed consent as an exercise of rights rather than a waiver of rights. Also, consent is not a waiver of all rights. For example, a patient does not waive his or her right to sue a physician who negligently provides a treatment harmful to the patient. In a truly informed consent, it should be clearly stated which rights, if any, are waived. See Manson and O'Neill, *Rethinking Informed Consent in Bioethics*, esp. pp. 72–77, 187–89. For a challenge to Manson and O'Neill's thesis, see Emma Bullock, "Informed Consent as Waiver: The Doctrine Rethought?" *Ethical Perspectives* 17 (2010): 529–55, available at http://www.ethical-perspectives.be/viewpic.php?LAN=E&TABLE=EP&ID=1277 (accessed May 8, 2018).
97. 97. On the last three examples, which we will not further pursue, see Alexander M. Capron, "Legal and Regulatory Standards of Informed Consent in Research," pp. 620–22.
98. 98. *Cobbs v. Grant*, 502 P.2d 1, 12 (1972).
99. 99. Baruch Brody, *Life and Death Decision Making* (New York: Oxford University Press, 1988), p. 22. The claim that rights to informed consent are always waivable is challenged in Rosemarie D. C. Bernabe et al., "Informed Consent and Phase IV Non-Interventional Drug Research," *Current Medical Research and Opinion* 27 (2011): 513–18.
100. 100. The Nuremberg Code, in *Trials of War Criminals before the Nuremberg Military Tribunals under Control Council Law no. 10* (Washington, DC: US Government Printing Office, 1949).
101. 101. See Joel Feinberg, *Social Philosophy* (Englewood Cliffs, NJ: Prentice Hall, 1973), p. 48; *Harm to Self*, pp. 112–18. For a notably different view of the concept of voluntariness and its connection to consent —one heavily influenced by law—see Paul S. Appelbaum, Charles W. Lidz, and Robert Klitzman, "Voluntariness of Consent to Research: A Conceptual Model," *Hastings Center Report* 39 (January–February 2009): 30–39, esp. 30–31, 33; and a criticism of Appelbaum, Lidz, and Klitzman in Robert M. Nelson, Tom L. Beauchamp, Victoria A. Miller, et al., "The Concept of Voluntary Consent," *American Journal of Bioethics* 11 (2011): 6–16, esp. 12–13.
102. 102. Our formulation is indebted to Robert Nozick, "Coercion," in *Philosophy, Science and Method: Essays in Honor of Ernest Nagel*, ed. Sidney Morgenbesser, Patrick Suppes, and Morton White (New York: St. Martin's, 1969), pp. 440–72; and Bernard Gert, "Coercion and Freedom," in *Coercion: Nomos XIV*, ed. J. Roland Pennock and John W. Chapman (Chicago: Aldine, Atherton, 1972), pp. 36–37. See in addition Alan Wertheimer, *Coercion* (Princeton, NJ: Princeton University Press, 1987).
103. 103. Cf. Jennifer S. Hawkins and Ezekiel J. Emanuel, "Clarifying Confusions about Coercion," *Hastings Center Report* 35 (September–October 2005): 16–19.
104. 104. For different views about the concept and ethics of manipulation, see Christian Coons and Michael Weber, eds., *Manipulation: Theory and Practice* (New York: Oxford University Press, 2014); Mark D. White, *The Manipulation of Choice: Ethics and Libertarian Paternalism* (New York: Palgrave Macmillan, 2013); Robert Noggle, "Manipulation, Salience, and Nudges," *Bioethics* 32, no. 3 (2018): 164–70; and Noggle, "The Ethics of Manipulation," *The Stanford Encyclopedia of Philosophy* (Summer 2018 Edition), ed. Edward N. Zalta, available at https://plato.stanford.edu/archives/sum2018/entries/ethics-manipulation/ (accessed October 8, 2018).

105. 105. See James H. Jones, *Bad Blood*, rev. ed. (New York: Free Press, 1993); David J. Rothman, "Were Tuskegee & Willowbrook 'Studies in Nature'?" *Hastings Center Report* 12 (April 1982): 5–7; Susan M. Reverby, ed., *Tuskegee's Truths: Rethinking the Tuskegee Syphilis Study* (Chapel Hill: University of North Carolina Press, 2000); Reverby, *Examining Tuskegee: The Infamous Syphilis Study and Its Legacy* (Chapel Hill: University of North Carolina Press, 2009); and Ralph V. Katz and Rueben Warren, eds., *The Search for the Legacy of the USPHS Syphilis Study at Tuskegee: Reflective Essays Based upon Findings from the Tuskegee Legacy Project* (Lanham, MD: Lexington Books, 2011).
106. 106. See Sarah E. Hewlett, "Is Consent to Participate in Research Voluntary," *Arthritis Care and Research* 9 (1996): 400–404; Victoria Miller et al., "Challenges in Measuring a New Construct: Perception of Voluntariness for Research and Treatment Decision Making," *Journal of Empirical Research on Human Research Ethics* 4 (2009): 21–31; and Nancy E. Kass et al., "Trust: The Fragile Foundation of Contemporary Biomedical Research," *Hastings Center Report* 26 (September–October 1996): 25–29.
107. 107. See Charles W. Lidz et al., *Informed Consent: A Study of Decision Making in Psychiatry* (New York: Guilford, 1984), chap. 7, esp. pp. 110–11, 117–23.
108. 108. U.S. federal regulations for research involving human subjects require "additional safeguards … to protect the rights and welfare" of subjects "likely to be vulnerable to coercion or undue influence, such as children, prisoners, individuals with impaired decision-making capacity, or economically or educationally disadvantaged persons," but the key concepts are inadequately analyzed and the list of groups is not uncontroversial. See Code of Federal Regulations, title 45, Public Welfare, Department of Health and Human Services, Part 46, Protection of Human Subjects, Subpart A ("Common Rule"), as revised in 2017 with general implementation January 21, 2019. For examinations of possible types of vulnerability in research involving human subjects, see Kenneth Kipnis, "Vulnerability in Research Subjects: A Bioethical Taxonomy," in National Bioethics Advisory Commission, *Ethical and Policy Issues in Research Involving Human Participants*, vol. 2 (Bethesda, MD: National Bioethics Advisory Commission, 2001): G1–13; and James DuBois, "Vulnerability in Research," in *Institutional Review Board: Management and Function*, 2nd ed., ed. Robert Amdur and Elizabeth Bankert (Boston: Jones & Bartlett, 2005), pp. 337–40.
109. 109. For the distinction between decisional autonomy and executional autonomy, see Bart J. Collopy, "Autonomy in Long Term Care," *Gerontologist* 28, Supplementary Issue (June 1988): 10–17. On failures to appreciate both capacity and incapacity, see C. Dennis Barton et al., "Clinicians' Judgement of Capacity of Nursing Home Patients to Give Informed Consent," *Psychiatric Services* 47 (1996): 956–60; and Meghan B. Gerety et al., "Medical Treatment Preferences of Nursing Home Residents," *Journal of the American Geriatrics Society* 41 (1993): 953–60.
110. 110. *Superintendent of Belchertown State School v. Saikewicz*, Mass. 370 N.E. 2d 417 (1977).
111. 111. For a survey of research on substituted judgment, see Daniel P. Sulmasy, "Research in Medical Ethics: Scholarship in 'Substituted Judgment,'" in *Methods in Medical Ethics*, 2nd ed., ed. Jeremy Sugarman and Daniel P. Sulmasy (Washington, DC: Georgetown University Press, 2010), pp. 295–314. For recent debates about conceptions and implementation of substituted judgment, see several articles in the *Journal of Medical Ethics* 41 (September 2015).
112. 112. See Rohit Devnani, James E. Slaven, Jr., Gabriel T. Bosslet, et al., "How Surrogates Decide: A Secondary Data Analysis of Decision-Making Principles Used by the Surrogates of Hospitalized Older Adults," *Journal of General Internal Medicine* 32 (2017): 1285–93.
113. 113. See, for example, *In the Matter of the Application of John Evans against Bellevue Hospital*, Supreme Court of the State of New York, Index No. 16536/87 (1987).
114. 114. A. D. Firlik, "Margo's Logo" (Letter), *JAMA: Journal of the American Medical Association* 265 (1991): 201.
115. 115. Ronald Dworkin, *Life's Dominion: An Argument about Abortion, Euthanasia, and Individual Freedom* (New York: Knopf, 1993), pp. 221–29.
116. 116. President's Council on Bioethics, *Taking Care: Ethical Caregiving in Our Aging Society* (Washington, DC: President's Council on Bioethics, September 2005), p. 84. The President's Council draws in part on the work of one of its members, Rebecca Dresser, "Dworkin on Dementia: Elegant Theory, Questionable Policy," *Hastings Center Report* 25 (November–December 1995): 32–38.

5

Nonmaleficence

The principle of nonmaleficence obligates us to abstain from causing harm to others. In medical ethics this principle has often been treated as effectively identical to the celebrated maxim *Primum non nocere*: "Above all [or first] do no harm." Often proclaimed the fundamental principle in the Hippocratic tradition, this principle does not appear in the Hippocratic writings, and a venerable statement sometimes confused with it—"at least, do no harm"—is a strained translation of a single Hippocratic passage.[1] Nonetheless, the Hippocratic oath incorporates both an obligation of nonmaleficence and an obligation of beneficence: "I will use treatment to help the sick according to my ability and judgment, but I will never use it to injure or wrong them."

This chapter explores the principle of nonmaleficence and its implications for several areas of biomedical ethics where harms may occur. We examine distinctions between killing and allowing to die, intending and foreseeing harmful outcomes, withholding and withdrawing life-sustaining treatments, as well as controversies about the permissibility of physicians assisting seriously ill patients in bringing about their deaths. The terminally ill and the critically ill and injured are featured in many of these discussions. The framework for decision making about life-sustaining procedures and assistance in dying that we defend would alter certain central features in traditional medical practice for both competent and incompetent patients. Central to our framework is a commitment to, rather than suppression of, quality-of-life judgments. This chapter also addresses moral problems in the protection of incompetent patients through advance directives and surrogate decision makers as well as special issues in decision making about children. Finally, the chapter examines the underprotection and the overprotection of subjects of research through public and institutional policies; and we also examine harms that can befall individuals and groups from unduly broad forms of consent in research on stored biological samples.

The Concept and Principle of Nonmaleficence

The Distinction between the Principles of Nonmaleficence and Beneficence

Many ethical theories recognize a principle of nonmaleficence.[2] Some philosophers combine nonmaleficence with beneficence to form a single principle. William Frankena, for example, divides the principle of beneficence into four general obligations, the first of which we identify as the principle and obligation of nonmaleficence and the other three of which we refer to as principles and obligations of beneficence:

1. 1. One ought not to inflict evil or harm.
2. 2. One ought to prevent evil or harm.
3. 3. One ought to remove evil or harm.
4. 4. One ought to do or promote good.[3]

If we were to bring these ideas of benefiting others and not injuring them under a single principle, we would be forced to note, as did Frankena, the several distinct obligations embedded in this general principle. In our view, conflating nonmaleficence and beneficence into a single principle obscures critical moral distinctions as well as different types of moral theory. Obligations not to harm others, such as those prohibiting theft, disabling, and killing, are distinct from obligations to help others, such as those prescribing the provision of benefits, protection of interests, and promotion of welfare.

Obligations not to harm others are sometimes more stringent than obligations to help them, but the reverse is also true. If in a particular case a health care provider inflicts a minor injury—swelling from a needlestick, say—but simultaneously provides a major benefit such as saving the patient's life, it is justified to conclude that the obligation of beneficence takes priority over the obligation of nonmaleficence in this case.[4] In many situations,

inflicting surgical harm to improve a patient's chance of survival, introducing social burdens to protect the public's health, and subjecting some research subjects to risks to generate valuable knowledge can all be justified by the intended benefits.

One might try to reformulate the common (but ultimately flawed) idea of nonmaleficence's increased stringency as follows: Obligations of nonmaleficence are usually more stringent than obligations of beneficence, and nonmaleficence may override beneficence, even if the best utilitarian outcome would be obtained by acting beneficently. If a surgeon, for example, could save two innocent lives by killing a prisoner on death row to retrieve his heart and liver for transplantation, this outcome of saving two lives would have the highest net utility under the circumstances, but the surgeon's action would be morally indefensible.

This formulation of stringency with respect to nonmaleficence has an initial ring of plausibility, but we need to be especially cautious about constructing axioms of priority. Nonmaleficence does sometimes override other principles, but the weights of these moral principles vary in different circumstances. In our view, no rule in ethics favors avoiding harm over providing benefit in every circumstance, and claims that an order of priority exists among elements 1 through 4 in Frankena's scheme is unsustainable.

Rather than attempting to structure a hierarchical ordering, we group the principles of nonmaleficence and beneficence into four norms that do not have an a priori rank order:

Nonmaleficence

1. 1. One ought not to inflict evil or harm.

Beneficence

1. 2. One ought to prevent evil or harm.
2. 3. One ought to remove evil or harm.
3. 4. One ought to do or promote good.

Each of the three principles of beneficence requires taking action by *helping*—preventing harm, removing harm, and promoting good—whereas nonmaleficence requires only *intentional avoidance* of actions that cause harm. Rules of nonmaleficence therefore take the form "Do not do X." Some philosophers accept only principles or rules that take this proscriptive form. They even limit rules of respect for autonomy to rules of the form "Do not interfere with a person's autonomous choices." These philosophers reject all principles or rules that require helping, assisting, or rescuing other persons, although they recognize these norms as legitimate *moral ideals*.[5] However, the mainstream of moral philosophy does not accept this sharp distinction between moral *obligations* of refraining and moral *ideals* of helping. Instead, it recognizes and preserves the relevant distinctions by distinguishing obligations of refraining from inflicting harm and obligations of helping. We take the same view, and in Chapter 6 (pp. 218–24), we explain further the nature of the distinction.

Legitimate disagreements arise about how to classify actions under categories 1 through 4 as well as about the nature and stringency of the obligations that arise from them. Consider the following case: Robert McFall was dying of aplastic anemia, and his physicians recommended a bone marrow transplant from a genetically compatible donor to increase his chances of living one additional year from 25% to a range of 40% to 60%. The patient's cousin, David Shimp, agreed to undergo tests to determine his suitability as a donor. After completing the test for tissue compatibility, he refused to undergo the test for genetic compatibility. He had changed his mind about donation. Robert McFall's lawyer asked a court to compel Shimp to undergo the second test and donate his bone marrow if the test indicated a good match.[6]

Public discussion focused on whether Shimp had an obligation of beneficence toward McFall in the form of an obligation to prevent harm, to remove harm, or to promote McFall's welfare. Though ultimately unsuccessful, McFall's lawyer contended that even if Shimp did not have a legal obligation of beneficence to rescue his cousin, he did have a legal obligation of nonmaleficence, which required that he not make McFall's situation worse. The lawyer argued that when Shimp agreed to undergo the first test and then backed out, he caused a

"delay of critical proportions" that constituted a violation of the obligation of nonmaleficence. The judge ruled that Shimp did not violate any legal obligation but also held that his actions were "morally indefensible."[7]

This case illustrates difficulties of identifying specific obligations under the principles of beneficence and nonmaleficence and shows the importance of *specifying* these principles (as discussed in our Chapters 1 and 10) to handle circumstances such as those of donating organs or tissues, withholding life-sustaining treatments, hastening the death of a dying patient, and biomedical research involving both human and animal subjects.

The Concept of Harm

The concept of nonmaleficence has been explicated by the concepts of *harm* and *injury*, but we will confine our analysis to harm. This term has both a normative and a nonnormative use. "X harmed Y" sometimes means that X wronged Y or treated Y unjustly, but it sometimes only means that X's action had an adverse effect on Y's interests. As we use these notions, *wronging* involves violating someone's rights, but *harming* need not signify such a violation. People are harmed without being wronged through attacks by disease, natural disasters, bad luck, and acts of others to which the harmed person has consented.[8] People can also be wronged without being harmed. For example, if an insurance company improperly refuses to pay a patient's hospital bill and the hospital shoulders the full bill, the insurance company wrongs the patient without harming him or her.

We construe harm as follows: A harm is a thwarting, defeating, or setting back of some party's interests, but a harmful action is not always wrong or unjustified.[9] Harmful actions that involve justifiable setbacks to another's interests are not wrong—for example, justified amputation of a consenting patient's leg, justified punishment of physicians for incompetence or negligence, justified demotion of employees for poor performance, and some forms of research involving animals. Nevertheless, the principle of nonmaleficence is a prima facie principle that requires the justification of harmful actions. This justification may come from showing that the harmful actions do not infringe specific obligations of nonmaleficence or that the infringements are outweighed by other ethical principles and rules.

Some definitions of *harm* are so broad that they include setbacks to interests in reputation, property, privacy, and liberty or, in some writings, discomfort, humiliation, and annoyance. Such broad conceptions can still distinguish trivial harms from serious harms by the magnitude of the interests affected. Other accounts with a narrower focus view harms exclusively as setbacks to physical and psychological interests, such as those in health and survival.

Whether a broad or a narrow construal is preferable is not a matter we need to decide here. Although harm is a contested concept, significant bodily harms and setbacks to other significant interests are paradigm instances of harm. We concentrate on physical and mental harms, especially pain, disability, suffering, and death, while recognizing other setbacks to interests. Intending, causing, and permitting death or the risk of death are especially important subjects.

Rules Specifying the Principle of Nonmaleficence

The principle of nonmaleficence supports several more specific moral rules (although principles other than nonmaleficence help justify some of these rules).[10] Examples of more specific rules include the following:[11]

1. 1. Do not kill.
2. 2. Do not cause pain or suffering.
3. 3. Do not incapacitate.
4. 4. Do not cause offense.
5. 5. Do not deprive others of the goods of life.

Both the principle of nonmaleficence and its specifications into these moral rules are prima facie binding, not absolute.

Negligence and the Standard of Due Care

Obligations of nonmaleficence include not only obligations not to inflict harms, but also obligations not to impose *risks* of harm. A person can harm or place another person at risk without malicious or harmful intent, and the agent of harm may or may not be morally or legally responsible for the harms. In some cases agents are causally responsible for a harm that they did not intend or know about. For example, if cancer rates are elevated at a chemical plant as the result of exposure to a chemical not previously suspected as a carcinogen, the employer has placed its workers at risk by its decisions or actions, even though the employer did not intentionally or knowingly cause the harm.

In cases of risk imposition, both law and morality recognize a standard of *due care* that determines whether the agent who is causally responsible for the risk is legally or morally responsible as well. This standard is a specification of the principle of nonmaleficence. Due care is taking appropriate care to avoid causing harm, as the circumstances demand of a reasonable and prudent person. This standard requires that the goals pursued justify the risks that must be imposed to achieve those goals. Grave risks require commensurately momentous goals for their justification. Serious emergencies justify risks that many nonemergency situations do not justify. For example, attempting to save lives after a major accident justifies, within limits, dangers created by rapidly moving emergency vehicles. A person who takes due care in this context does not violate moral or legal rules even if significant risk for other parties is inherent in the attempted rescue.

Negligence falls short of due care. In professions, negligence involves a departure from the professional standards that determine due care in given circumstances. The term *negligence* covers two types of situations: (1) intentionally imposing unreasonable risks of harm (advertent negligence or recklessness) and (2) unintentionally but carelessly imposing risks of harm (inadvertent negligence). In the first type, an agent knowingly imposes an unwarranted risk: For example, a nurse knowingly fails to change a bandage as scheduled, creating an increased risk of infection. In the second type, an agent unknowingly performs a harmful act that he or she should have known to avoid: For example, a physician acts negligently if he or she knows but forgets that a patient does not want to receive certain types of information and discloses that information, causing fear and shame in the patient. Both types of negligence are morally blameworthy, although some conditions may mitigate blameworthiness.[12]

In treating negligence, we will concentrate on conduct that falls below a standard of due care that law or morality establishes to protect others from the careless imposition of risks. Courts must determine responsibility and liability for harm, when a patient, client, or consumer seeks compensation for setbacks to interests or punishment of a responsible party, or both. We will not concentrate on legal liability and instead will adapt parts of the legal model of responsibility for harmful action to formulate moral responsibility for harm caused by health care professionals. The following are essential elements in this professional model of due care:

1. 1. The professional must have a duty to the affected party.
2. 2. The professional must breach that duty.
3. 3. The affected party must experience a harm.
4. 4. The harm must be caused by the breach of duty.

Professional malpractice is an instance of negligence that involves failure to follow professional standards of care.[13] By entering into the profession of medicine, physicians accept a responsibility to observe the standards specific to their profession. When a therapeutic relationship proves harmful or unhelpful, malpractice occurs if and only if physicians do not meet professional standards of care. For example, in *Adkins v. Ropp* the Supreme Court of Indiana considered a patient's claim that a physician acted negligently in removing foreign matter from the patient's eye:

> When a physician and surgeon assumes to treat and care for a patient, in the absence of a special agreement, he is held in law to have impliedly contracted that he possesses the reasonable and ordinary qualifications of his profession and that he will exercise at least reasonable skill, care, and diligence in his treatment of him. This implied contract on the part of the physician does not include

> a promise to effect a cure and negligence cannot be imputed because a cure is not effected, but he does impliedly promise that he will use due diligence and ordinary skill in his treatment of the patient so that a cure may follow such care and skill. This degree of care and skill is required of him, not only in performing an operation or administering first treatments, but he is held to the same degree of care and skill in the necessary subsequent treatments unless he is excused from further service by the patient himself, or the physician or surgeon upon due notice refuses to further treat the case.[14]

The line between due care and inadequate care is sometimes difficult to draw. Increased safety measures in epidemiological and toxicological studies, educational and health promotional programs, and other training programs can sometimes reduce health risks. However, a substantial question remains about the lengths to which physicians, employers, and others must go to avoid or to lower risks—a moral problem of determining the scope of obligations of nonmaleficence.

DISTINCTIONS AND RULES GOVERNING NONTREATMENT DECISIONS

Religious traditions, philosophical discourse, professional codes, public policy, and law have developed many guidelines to specify the requirements of nonmaleficence in health care, particularly with regard to treatment and nontreatment decisions. Some of these guidelines are helpful, but others need revision or replacement. Many draw heavily on at least one of the following distinctions:

1. 1. *Withholding* and *withdrawing* life-sustaining treatment
2. 2. *Medical treatments* and *artificial nutrition and hydration*
3. 3. *Intended* effects and *merely foreseen* effects

Although at times influential in medicine and law, these distinctions, we will argue, are outmoded and need to be replaced. The venerable position that these traditional distinctions have occupied in professional codes, institutional policies, and writings in biomedical ethics by itself provides no warrant for retaining them when they are obsolete, no longer helpful, and sometimes even morally dangerous.

Withholding and Withdrawing Treatments

Debate about the principle of nonmaleficence and forgoing life-sustaining treatments has centered on the omission-commission distinction, especially the distinction between withholding (not starting) and withdrawing (stopping) treatments. Many professionals and family members feel justified in withholding treatments they never started, but not in withdrawing treatments already initiated. They sense that decisions to stop treatments are more momentous, consequential, and morally fraught than decisions not to start them. Stopping a respirator, for example, seems to many to cause a person's death, whereas not starting the respirator does not seem to have this same causal role.[15]

In one case, an elderly man suffered from several major medical problems with no reasonable chance of recovery. He was comatose and unable to communicate. Antibiotics to fight infection and an intravenous (IV) line to provide nutrition and hydration kept him alive. No evidence indicated that he had expressed his wishes about life-sustaining treatments while competent, and he had no family member to serve as a surrogate decision maker. Physicians and staff quickly agreed on a "no code" or "do not resuscitate" (DNR) order, a signed order not to attempt cardiopulmonary resuscitation if a cardiac or respiratory arrest occurred. In the event of such an arrest, the patient would be allowed to die. The staff felt comfortable with this decision because of the patient's overall condition and prognosis, and because they could view not resuscitating the patient as withholding rather than withdrawing treatment.

Questions arose about whether to continue the interventions in place. Some members of the health care team thought that they should stop all medical treatments, including antibiotics and artificial nutrition and hydration, because, in their language, these treatments were "extraordinary" or "heroic."[16] Others thought it wrong to stop

these treatments once they had been started. A disagreement erupted about whether it would be permissible not to insert the IV line again if it became infiltrated—that is, if it broke through the blood vessel and began leaking fluid into surrounding tissue. Some who had opposed stopping treatments were comfortable with not inserting the IV line again, because they viewed the action as withholding rather than withdrawing. They emphatically opposed reinsertion if it required a cutdown (an incision to gain access to the deep large blood vessels) or a central line. Others viewed the provision of artificial nutrition and hydration as a single process and felt that inserting the IV line again was simply continuing what had been interrupted. For them, not restarting was equivalent to withdrawing and thus, unlike withholding, morally wrong.[17]

In many similar cases caregivers' discomfort about withdrawing life-sustaining treatments appears to reflect the view that such actions render them causally responsible and morally or legally culpable for a patient's death, whereas they are not responsible if they never initiate a life-sustaining treatment. The conviction that starting a treatment often creates valid claims or expectations for its continuation is another source of caregiver discomfort. Only if patients waive the claim for continued treatment does it seem legitimate to many caregivers to stop procedures. Otherwise, stopping procedures appears to breach expectations, promises, or contractual obligations to the patient, family, or surrogate decision maker. Patients for whom physicians have not initiated treatment seem to hold no parallel claim.[18]

Feelings of reluctance about withdrawing treatments are understandable, but the distinction between withdrawing and withholding treatments is morally irrelevant and potentially dangerous. The distinction is unclear, inasmuch as withdrawing can happen through an omission (withholding) such as not recharging batteries that power respirators or not putting the infusion into a feeding tube. In multi-staged treatments, decisions not to start the next stage of a treatment plan can be tantamount to stopping treatment, even if the early phases of the treatment continue.

Both not starting and stopping can be justified, depending on the circumstances. Both can be instances of allowing to die, and both can be instances of killing. Courts recognize that individuals can commit a crime by omission if they have an obligation to act, just as physicians can commit a wrong by omission in medical practice. Such judgments depend on whether a physician has an obligation either not to withhold or not to withdraw treatment. In these cases if a physician has a duty to treat, omission of treatment breaches this duty, whether or not withholding or withdrawing is involved. However, if a physician does not have a duty to treat or has a duty not to treat, omission of either type involves no moral violation. Indeed, if the physician has a duty not to treat, it would be morally wrong to start the treatment or to continue the treatment if it has already begun.

In a classic case (to be discussed further later in this chapter), a court raised the following legal problem about continuing kidney dialysis for Earle Spring, an elderly patient with numerous medical problems: "The question presented by … modern technology is, once undertaken, at what point does it cease to perform its intended function?" The court held that "a physician has no duty to continue treatment, once it has proven to be ineffective." The court emphasized the need to balance benefits and burdens to determine overall effectiveness.[19] Although legal responsibility cannot be equated with moral responsibility in such cases, the court's conclusion is consistent with the moral conclusions about justified withdrawal for which we are presently arguing. Approximately one in four deaths of patients with end-stage renal disease in the United States occurs after a decision to withdraw dialysis.[20] The practice is common, and the decisions are often justified.[21]

Giving priority to withholding over withdrawing treatment can lead to *overtreatment* in some cases, that is, the continuation of no longer beneficial or desirable treatment for the patient. Less obviously, the distinction can lead to *undertreatment*. Patients and families may worry about being trapped by biomedical technology that, once begun, cannot be stopped. To circumvent this problem, they may become reluctant to authorize the technology even when it could possibly benefit the patient. Health care professionals sometimes display the same reluctance. In one case, a seriously ill newborn died after several months of treatment, much of it against the parents' wishes, because a physician was unwilling to stop the respirator once it had been connected. Later this physician reportedly felt "less eager to attach babies to respirators now."[22]

The moral burden of proof is often heavier when the decision is to withhold rather than to withdraw treatments. Only after starting treatments will it be possible, in many cases, to make a proper diagnosis and prognosis as well as to balance prospective benefits and burdens. This trial period can reduce uncertainty about outcomes. Patients and surrogates often feel less stress and more in control if they can reverse or otherwise change a decision to treat after the treatment has started. Accordingly, responsible health care may involve a trial with periodic reevaluation. Caregivers then have time to judge the effectiveness of the treatment, and the patient or surrogate has time to evaluate its benefits and burdens. Not to propose or allow the trial is morally worse than not trying, and withholding may be worse than withdrawing in these cases.

To a great extent, the withholding-withdrawing distinction has shaped an intense debate about cardiovascular implantable electronic devices (CIEDs), which include pacemakers and implantable cardioverter-defibrillators (ICDs). These devices are increasingly common and often helpful and necessary. While clinicians have generally been comfortable in not implanting these devices when patients or their surrogates do not want them, they have often been uncomfortable discontinuing them, particularly pacemakers, even though each one can be stopped noninvasively, without surgery. Horror stories abound. In one case, a woman described the struggle to have her elderly, severely demented, significantly incapacitated father's battery-powered pacemaker turned off. The pacemaker had been inserted because, without it, a cardiologist would not clear her father for surgery to correct a painful intestinal hernia. The family later realized that a temporary version would have sufficed. When her father's health problems worsened, and her mother requested deactivation of the pacemaker, the physician refused because "it would have been like putting a pillow over [his] head."[23]

Many physicians, over 60% in one study,[24] see an ethical distinction between deactivating a pacemaker and deactivating an ICD. For some, deactivation of pacemakers is tantamount to active euthanasia. This morally dubious judgment is rooted in the fact that pacemakers provide continuous rather than intermittent treatment and their removal may lead to immediate death, thereby increasing the professional's sense of causal and moral responsibility.[25] A consensus statement in 2010, involving several professional groups, rightly dismissed any ethical and legal distinctions among CIEDs, viewing all of them as life-sustaining treatments that patients and their surrogates may legitimately request to be withdrawn in order to allow the underlying disease to take its course.[26] The consensus statement recognized clinicians' rights not to participate in the withdrawal while, at the same time, emphasizing their responsibility to refer patients to clinicians or others who would deactivate the devices. As it happens, industry representatives deactivate the pacemaker about half the time and the ICD about 60% of the time.[27]

We conclude that the distinction between withholding and withdrawing is morally untenable and can be morally dangerous. If a clinician makes decisions about treatment using this irrelevant distinction, or allows a surrogate (without efforts at dissuasion) to make such a decision, the clinician is morally blameworthy for negative outcomes. The felt importance of the distinction between not starting and stopping procedures undoubtedly accounts for, but does not justify, the speed and ease with which hospitals and health care professionals decades ago accepted no code or DNR orders and formed hospital policies regarding cardiopulmonary resuscitation (CPR). Policies regarding CPR often stand independent of other policies governing life-sustaining technologies, such as respirators, in part because many health care professionals view not providing CPR as withholding rather than withdrawing treatment. Clinicians' decisions to withhold CPR, through "do-not-attempt resuscitation" (DNAR) or "do-not-resuscitate" (DNR) orders, are ethically problematic when made unilaterally without advance consultation with patients and/or their families or, generally but not always, against their requests.[28] (See further our discussion of futile interventions below and in Chapter 6.)

Medical Treatments and Artificial Nutrition and Hydration

Widespread debate has occurred about whether the distinction between *medical* technologies and *artificial nutrition and hydration* (AN&H), which might be called *sustenance* technologies, can be used to differentiate between justified and unjustified forgoing of life-sustaining treatments. Some argue that technologies for supplying nutrition and hydration using needles, tubes, catheters, and the like should be sharply distinguished from medical life-sustaining technologies, such as respirators and dialysis machines. Others dispute this

distinction, contending that the technologies for AN&H are relevantly similar to other medical technologies[29] and therefore should be subject to the same framework of ethical analysis and assessment.[30]

To help determine whether this distinction is defensible and useful, we examine some cases, beginning with the case of a seventy-nine-year-old widow who had resided in a nursing home for several years, frequently visited by her daughter and grandchildren, who loved her deeply. In the past she experienced repeated transient ischemic attacks caused by reductions or stoppages of blood flow to the brain. Because of progressive organic brain syndrome, she had lost most of her mental abilities and had become disoriented. She also had thrombophlebitis (inflammation of a vein associated with clotting) and congestive heart failure. One day she suffered a massive stroke. She made no recovery, remained nonverbal, manifested a withdrawal reaction to painful stimuli, and exhibited a limited range of purposeful behaviors. She strongly resisted a nasogastric tube being placed into her stomach to introduce nutritional formulas and water. At each attempt she thrashed about violently and pushed the tube away. When the tube was finally placed, she managed to remove it. After several days the staff could not find new sites for inserting IV lines, and debated whether to take further measures to maintain fluid and nutritional intake for this elderly patient, who did not improve and was largely unaware and unresponsive. After lengthy discussions with nurses on the floor and with the patient's family, the physicians in charge concluded that they should not provide further IVs, cutdowns, or a feeding tube. The patient had minimal oral intake and died quietly the following week.[31]

Second, in a groundbreaking case in 1976, the New Jersey Supreme Court ruled it permissible for a guardian to disconnect Karen Ann Quinlan's respirator and allow her to die.[32] After the respirator was removed, Quinlan lived for almost ten years, protected by antibiotics and sustained by nutrition and hydration provided through a nasogastric tube. Unable to communicate, she lay comatose in a fetal position, with increasing respiratory problems, bedsores, and weight loss from 115 to 70 pounds. A moral issue developed over those ten years. If it is permissible to remove the respirator, is it permissible to remove the feeding tube? Several Roman Catholic moral theologians advised the parents that they were not morally required to continue medically administered nutrition and hydration or antibiotics to fight infections. Nevertheless, the Quinlans continued AN&H because they believed that the feeding tube did not cause pain, whereas the respirator did.[33]

US courts have since generally placed AN&H under the same substantive and procedural standards as other medical treatments such as the respirator.[34] In the much-discussed Terri Schiavo case, the husband and parents of a woman who was in a persistent vegetative state (PVS) were in conflict over whether it was justifiable to withdraw her feeding tube. Despite legal challenges and political conflicts, the court applying Florida's laws allowed the husband, expressing what he represented as Terri Schiavo's wishes, to withdraw AN&H to allow her to die, approximately fifteen years after she entered the PVS.[35]

It is understandable that some familial and professional caregivers find cultural, religious, symbolic, or emotional barriers to withholding or withdrawing AN&H from patients.[36] They sometimes describe withholding or withdrawing AN&H as "starving" or letting a patient "starve" to death.[37] And some state laws and public and institutional policies also express this sentiment, particularly for patients in PVS. However, in our judgment, caregivers may justifiably forgo AN&H for patients in some circumstances, as holds true for other life-sustaining technologies. No morally relevant difference exists between the various life-sustaining technologies, and the right to refuse medical treatment for oneself or others is not contingent on the type of treatment. There is no reason to believe that AN&H is always an essential part of palliative care or that it necessarily constitutes a beneficial medical treatment in all cases. Available evidence indicates that many terminally ill patients, including those with advanced dementia, die more comfortably without AN&H, which, of course, should always be provided when needed for comfort.[38]

Intended Effects and Merely Foreseen Effects: The Rule of Double Effect

Another venerable attempt to specify the principle of nonmaleficence appears in the rule of double effect (RDE), also called the principle or doctrine of double effect. This rule incorporates an influential distinction between

intended effects and merely foreseen effects.

Functions and conditions of the RDE. The RDE is invoked to justify claims that a single act, which has one or more good effects and one or more harmful effects (such as death), is not always morally prohibited.[39] As an example of the use of the RDE, consider a patient experiencing terrible pain and suffering who asks a physician for help in ending his life. Suppose the physician injects the patient with a chemical to intentionally cause the patient's death as a means to end the patient's pain and suffering. The physician's action is wrong, under the RDE, because it involves the intention and direct means to cause the patient's death. In contrast, suppose the physician could provide medication to relieve the patient's pain and suffering at a substantial risk that the patient would die as a result of the medication. If the physician refuses to administer the medication, the patient will endure continuing pain and suffering; if the physician provides the medication, it may hasten the patient's death. If the physician intended, through the provision of medication, to relieve grave pain and suffering and did not intend to cause death, then the act of indirectly hastening death is not wrong, according to the mainline interpretation of the RDE.

Classical formulations of the RDE identify four conditions or elements that must be satisfied for an act with a double effect to be justified. Each is a necessary condition, and together they form sufficient conditions of morally permissible action:[40]

1. 1. *The nature of the act.* The act must be good, or at least morally neutral, independent of its consequences.
2. 2. *The agent's intention.* The agent intends only the good effect, not the bad effect. The bad effect can be foreseen, tolerated, and permitted, but it must not be intended.
3. 3. *The distinction between means and effects.* The bad effect must not be a means to the good effect. If the good effect were the causal result of the bad effect, the agent would intend the bad effect in his or her pursuit of the good effect.
4. 4. *Proportionality between the good effect and the bad effect.* The good effect must outweigh the bad effect. That is, the bad effect is permissible only if a proportionate reason compensates for permitting the foreseen bad effect.

All four conditions are controversial. We begin to investigate the cogency of the RDE by considering four cases of what many call therapeutic abortion (limited to protecting maternal life in these examples): (1) A pregnant woman has cervical cancer; she needs a hysterectomy to save her life, but this procedure will result in the death of the fetus. (2) A pregnant woman has an ectopic pregnancy—the nonviable fetus is in the fallopian tube—and physicians must remove the tube to prevent hemorrhage, which will result in the death of the fetus. (3) A pregnant woman has a serious heart disease that probably will result in her death if she attempts to carry the pregnancy to term. (4) A pregnant woman in difficult labor will die unless the physician performs a craniotomy (crushing the head of the unborn fetus). Some interpretations of Roman Catholic teachings, where the RDE has been prominent, hold that the actions that produce fetal death in the first two cases sometimes satisfy the four conditions of the RDE and therefore can be morally acceptable, whereas the actions that produce fetal death in the latter two cases never meet the conditions of the RDE and therefore are always morally unacceptable.[41]

In the first two cases, according to proponents of the RDE, a physician undertakes a legitimate medical procedure aimed at saving the pregnant woman's life with the foreseen but unintended result of fetal death. When viewed as unintended side effects (rather than as ends or means), these fetal deaths are said to be justified by the proportionately grave reason of saving the pregnant woman's life. In both of the latter two cases, the action of terminating fetal life is a *means* to save the pregnant woman's life. As such, it requires intending the fetus's death even if the death is not desired. Therefore, in those cases, criteria 2 and 3 are violated and the act cannot be justified by proportionality (criterion 4).

However, it is not likely that a morally relevant difference can be established between cases such as a hysterectomy or a craniotomy in terms of the abstract conditions that comprise the RDE. In neither case does the agent want or desire the death of the fetus, and the descriptions of the acts in these cases do not indicate morally relevant differences between intending, on the one hand, and foreseeing but not intending, on the other. It is

unclear why advocates of RDE conceptualize craniotomy as killing the fetus rather than as the act of crushing the skull of the fetus with the unintended result that the fetus dies. Similarly, it remains unclear why in the hysterectomy case the death of the fetus is foreseen but not intended. Proponents of the RDE must have a practicable way to distinguish the intended from the merely foreseen, but they face major difficulties in providing a theory of intention precise enough to draw defensible moral lines between the hysterectomy and craniotomy cases.

A problematic conception of intention. Adherents of the RDE need an account of intentional actions and intended effects of action to distinguish them from nonintentional actions and unintended effects. The literature on intentional action is itself controversial and focuses on diverse conditions such as volition, deliberation, willing, reasoning, and planning. One of the few widely shared views in this literature is that intentional actions require that an agent have a plan—a blueprint, map, or representation of the means and ends proposed for the execution of an action.[42] For an action to be intentional, it must correspond to the agent's plan for its performance.

Alvin Goldman uses the following example in an attempt to prove that agents do not intend merely foreseen effects.[43] Imagine that Mr. G takes a driver's test to prove competence. He comes to an intersection that requires a right turn and extends his arm to signal for a turn, although he knows it is raining and that he will get his hand wet. According to Goldman, Mr. G's signaling for a turn is an intentional act. By contrast, his getting a wet hand is an unintended effect or "incidental by-product" of his hand-signaling. A defender of the RDE must elect a similarly narrow conception of what is intended to avoid the conclusion that an agent intentionally brings about all the consequences of an action that the agent foresees. The defender distinguishes between acts and effects, and then between effects that are desired or wanted and effects that are foreseen but not desired or wanted. The RDE views the latter effects as foreseen, but not intended.

It is better, we suggest, to discard the language of "wanting" and to say that foreseen, undesired effects are "tolerated."[44] These effects are not so undesirable that the actor would avoid performing the act that results in them; the actor includes them as a part of his or her plan of intentional action. To account for this point, we use a model of intentionality based on what is *willed* rather than what is *wanted*. On this model, intentional actions and intentional effects include any action and any effect specifically willed in accordance with a plan, including tolerated as well as wanted effects.[45] In this conception a physician can desire not to do what he intends to do, in the same way that one can be willing to do something but, at the same time, reluctant to do it or even detest doing it.

Under this conception of intentional acts and intended effects, the distinction between what agents intend and what they merely foresee in a planned action is not viable.[46] For example, if a man enters a room and flips a switch that he knows turns on both a light and a fan, but desires only to activate the light, he cannot say that he activates the fan unintentionally. Even if the fan made an obnoxious whirring sound that he is aware of and wants to avoid, it would be mistaken to say that he unintentionally brought about the obnoxious noise by flipping the switch. More generally, a person who knowingly and voluntarily acts to bring about an effect brings about that effect intentionally. The person intends the effect, but does not desire it, does not will it for its own sake, and does not intend it as the goal of the action.

The moral relevance of the RDE and its distinctions can be evaluated in light of this model of intention. Is it plausible to distinguish morally between intentionally causing the death of a fetus by craniotomy and intentionally removing a cancerous uterus that causes the death of a fetus? In both actions the intention is to save the woman's life with knowledge that the fetus will die as a result of the action. No agent in either scenario desires the negative result (the fetus's death) for its own sake, and none would have tolerated the negative result if its avoidance were morally preferable to the alternative outcome. All parties accept the bad effect only because they cannot eliminate it without sacrificing the good effect.

In the standard interpretation of the RDE, the fetus's death is a *means* to saving a woman's life in the unacceptable case but merely a *side effect* in the acceptable case. That is, an agent intends a means but does not intend a side effect. This approach seems to allow persons to foresee almost anything as a side effect rather than

as an intended means. It does not follow, however, that people can create or direct intentions as they please. For example, in the craniotomy case, the surgeon might not intend the death of the fetus but only intend to remove it from the birth canal. The fetus will die, but is this outcome more than an unwanted and, in double effect theory, unintended consequence?[47] We consider the outcome to be an unwanted but tolerated and intended consequence.

The RDE might appear to fare better in handling problems in the care of dying patients where there is no conflict between different parties. It is often invoked to justify a physician's administration of medication to relieve pain and suffering (the primary intention and effect) even when it will probably hasten the patient's death (the unintended, secondary effect). A related practice, terminal sedation, challenges the boundaries and use of the RDE. In terminal sedation, physicians induce a deep sleep or unconsciousness to relieve pain and suffering in the expectation that this state will continue until the patient dies. Some commentators contend that some cases of terminal sedation can be justified under the RDE, whereas others argue that terminal sedation directly, although slowly, kills the patient and thus is a form of euthanasia.[48] Much depends on the description of terminal sedation in a particular set of circumstances, including the patient's overall condition, the proximity of death, and the availability of alternative means to relieve pain and suffering, as well as the intention of the physician and other parties. Interpretations of the RDE to cover some cases of terminal sedation allow compassionate acts of relieving pain, suffering, and discomfort that will foreseeably hasten death.

Often in dispute is whether death is good or bad for a particular person, and nothing in the RDE settles this dispute. The RDE applies only in cases with both a bad and a good effect, but determining the goodness and badness of different effects is a separate judgment. Accordingly, the goodness or badness of death for a particular person, whether it occurs directly or indirectly, must be determined and defended on independent grounds.[49]

Defenders of the RDE eventually may solve the puzzles and problems that critics like us have identified, but they have not succeeded thus far. However, we suggest that one constructive effort to retain an emphasis on intention without entirely abandoning the larger point of the RDE would focus on the way actions display a person's motives and character.[50] In the case of performing a craniotomy to save a pregnant woman's life, a physician may not *want* or *desire* the death of the fetus and may regret performing a craniotomy just as much as he or she would in the case of removing a cancerous uterus. Such facts about the physician's motives and character can make a decisive difference to a moral assessment of the action and the agent, but this moral conclusion also can be reached independently of the RDE.

Optional Treatments and Obligatory Treatments

We have now rejected some common distinctions and rules about forgoing life-sustaining treatment and causing death that are accepted in some traditions of medical ethics. To replace them we propose a basic distinction between obligatory and optional treatments. Our replacement analysis relies heavily on quality-of-life considerations that are clearly incompatible with some of the distinctions and rules we have already rejected. The following categories are central to our arguments:

1. 1. Obligatory to Treat (Wrong Not to Treat)
2. 2. Obligatory Not to Treat (Wrong to Treat)
3. 3. Optional Whether to Treat (Neither Required nor Prohibited to Treat)

Under 3, the question is whether it is morally neutral and therefore optional to provide or not to provide a treatment.

The principles of nonmaleficence and beneficence have often been specified to establish a presumption in favor of providing life-sustaining treatments for sick and injured patients. However, this presumption has rarely been thought to entail that it is always obligatory to provide the treatments. The use of life-sustaining treatments sometimes violates patients' interests. For example, pain can be so severe and physical restraints so burdensome

that these factors outweigh anticipated benefits, such as brief prolongation of life. Providing the treatment may even be inhumane or cruel. In the case of some severely incompetent and suffering patients, the burdens can so outweigh the benefits that the treatment is wrong, not merely optional.

Conditions for Overriding the Prima Facie Obligation to Treat

Several conditions justify decisions by patients, surrogates, or health care professionals to withhold or withdraw treatment. We examine these conditions (in addition to valid refusal of treatment) in this section.

Futile or pointless interventions. Physicians have no obligation to provide pointless, futile, or contraindicated treatment. In an extreme example, if a patient has died but remains on a respirator, cessation of treatment cannot harm him or her, and a physician has no obligation to continue to treat. However, some religious and personal belief systems do not consider a patient dead according to the same criteria many health care institutions recognize. For example, if there is heart and lung function, even when only maintained by technology, some religious traditions hold that the person is not dead and that the treatment is not futile even if health care professionals deem it futile, useless, or wasteful. This example is the tip of an iceberg of controversies about futility.

Typically the term *futile* refers to a situation in which irreversibly dying patients have reached a point at which further treatment provides no medical benefit or is hopeless, and therefore is optional from a medical and moral point of view. Palliative interventions may and generally should be continued to relieve pain, suffering, and discomfort. This model of futility covers only some treatments that have been deemed futile. Less typically in the literature on futility all of the following have been labeled futile: (1) whatever will not produce a sought physiological effect (e.g., antibiotics for a viral infection), (2) whatever proposed intervention is completely speculative because it is an untried "treatment," (3) whatever is highly unlikely to have a good effect, (4) whatever probably will produce only a low-grade, insignificant outcome (i.e., the results are expected to be exceedingly poor), (5) whatever is highly likely to be more burdensome than beneficial, and (6) whatever—after balancing effectiveness, potential benefit, and potential risk or burden—warrants withdrawing or withholding treatment.[51] Accordingly, the term *futility* is used to cover various situations of improbable effects, improbable success, and unacceptable benefit-burden ratios. In our view, the first three and even the fourth could plausibly be labeled judgments of futility, while five and six are better understood as judgments of utility or proportionality, because they involving balancing benefits, burdens, and risks to the patient.

The plethora of competing conceptions and uncertain meanings in discussions of *futility* suggests that we should, wherever possible, avoid the term in favor of more precise language in deliberations and communications between the health care team and patients and families. Judgments of futility presuppose an accepted goal in relation to which an intervention is deemed to be useless. Because of a lack of consensus about "medical futility," the language of "inappropriate" or "potentially inappropriate" has gained traction and wider acceptance.[52] Recommendations by key organizations of critical care specialists in the United States and Europe have played a significant role in these changes.[53] An American statement proposes that the term "potentially inappropriate" be used in place of "futile" when interventions have at a minimum some chance of accomplishing the patient-sought goal "but clinicians believe that competing ethical considerations justify not providing them." This proposal does not altogether eliminate the term *futile*. Rather, its meaning and use are restricted narrowly to "the rare situations" in which patients or surrogates "request interventions that simply cannot accomplish their intended physiologic goal." In these situations, clinicians should not provide the futile interventions as a matter of good ethics and good clinical judgment.[54] This use of the term *futile* is narrower than ours, but that fact is less problematic than its invocation of the vague and unhelpful language of "inappropriate" to cover situations in which interventions can achieve some patient-sought goals but are outweighed by competing ethical considerations. Without greater clarity and precision, it is implausible to think that one can successfully describe, in deliberations or communications within a medical team or with a patient or family, what makes a particular intervention "inappropriate."[55] If the competing ethical considerations involve an unfavorable balance of probable benefits and probable burdens and harms to the patient, then this judgment needs to be articulated and defended with precision. It is not adequately captured by the nebulous language of "inappropriate" or

"potentially inappropriate." If one judges that these considerations involve competing claims of just and fair access to resources, that judgment also needs to be articulated and defended.

Ideally, in reaching judgments of futility in our sense, health care providers will focus on objective medical factors in their decisions involving the dead and the irreversibly dying. Realistically, however, this ideal is difficult to satisfy. Disagreement often exists among health professionals, and conflicts may arise from a family's belief in a possible miracle, a religious tradition's insistence on doing everything possible in such circumstances, and the like. It is sometimes difficult to know whether a judgment of futility is based on a probabilistic prediction of failure or on something closer to medical certainty. If an elderly patient has a 1% chance of surviving an arduous and painful regimen, one physician may call the procedure futile while another may view survival as unlikely, but a possibility meriting consideration. At stake is a value judgment about what is worth the effort, as well as scientific knowledge and evidence. The term *futility* typically expresses a combined value judgment (such as "the proposed intervention is useless relative to the goal that is sought") and scientific judgment (such as "available data show that …").

A physician is not morally required to provide a genuinely futile or contraindicated intervention and in some cases may be required *not* to provide the intervention. The physician may not even be required to mention an intervention that would be genuinely futile. These circumstances often involve incompetent patients, especially patients in a PVS, where physicians or hospital policies sometimes impose on patients or surrogates decisions to forgo life support. Hospitals are increasingly adopting policies aimed at denying interventions that physicians knowledgeably judge to be futile, especially after trying them for a reasonable period of time. However, the possibility of judgmental error by physicians should lead to caution in formulating these policies. Unreasonable demands by patients and families should not be given priority over reasonable policies and assessments in health care institutions. Respect for the autonomy of patients or authorized surrogates is not a trump that allows patients or families to determine, without medical assistance and agreement, that a treatment is or is not futile. The right to refuse a proposed intervention does not translate into a right to request or demand a particular intervention.

We conclude that a genuinely futile medical intervention—one that has no chance of being successful in light of acceptable medical goals—is morally optional and in many cases ought not be introduced or continued. However, undertaking a futile intervention, such as CPR, may be an act of compassion and care toward the grief-stricken family of a critically ill patient, and could be justified, within limits, to achieve a goal such as allowing time for additional family members to arrive to have a little time with the patient prior to death.[56] Legitimate disagreements about whether a medical intervention is futile in particular circumstances may be best resolved through institutional procedures such as mediation, ethics consultations, or ethics committee review, or, occasionally, judicial review.[57]

Burdens of treatment outweigh its benefits. Medical codes and institutional policies often mistakenly assume that physicians may legitimately terminate life-sustaining treatments for persons not able to consent to or refuse the treatments only if the patient is terminally ill. Even if the patient is not terminally ill, life-sustaining medical treatment is not obligatory if its overall burdens outweigh its benefits to the patient. Medical treatment for those not terminally ill is sometimes optional even if the treatment could prolong life indefinitely, the patient is incompetent, and no advance directive exists. Moral considerations of nonmaleficence do not demand the maintenance of biological life and do not require the initiation or continuation of treatment without regard to the patient's pain, suffering, and discomfort.

As an example, consider the case we mentioned earlier of seventy-eight-year-old Earle Spring who developed numerous medical problems, including chronic organic brain syndrome and kidney failure. Hemodialysis controlled the latter problem. Although several aspects of this case were never resolved—such as whether Spring was aware of his surroundings and able to express his wishes—a plausible argument existed that the family and health care professionals were not morally obligated to continue hemodialysis because of the balance of benefits and burdens to a patient whose compromised mental condition and kidney function would gradually worsen regardless of what was done. However, in this case, as in many others, a family conflict of interest

complicated the situation: The financially strapped family had to pay the mounting health care costs while attempting to make judgments in the patient's best interests.

We return later in this chapter in the subsection "Surrogate Decision Making without Advance Directives" to procedures designed to protect such incompetent patients in difficult situations.

Quality-of-Life Judgments

Controversies about quality-of-life judgments. Our arguments thus far give considerable weight to quality-of-life judgments in determining whether treatments are optional or obligatory. We have relied on the premise that when quality of life is sufficiently low and an intervention produces more harm than benefit for the patient, caregivers may justifiably withhold or withdraw treatment. However, these judgments require defensible criteria of benefits and burdens that avoid reducing quality-of-life judgments to arbitrary personal preferences or to the patient's social worth.

In a landmark legal and bioethics case involving quality-of-life judgments, sixty-eight-year-old Joseph Saikewicz, who had an IQ of 10 and a mental age of approximately two years and eight months, suffered from acute myeloblastic monocytic leukemia. Chemotherapy would have produced extensive suffering and possibly serious side effects. Remission under chemotherapy occurs in only 30% to 50% of such cases and typically only for two to thirteen months. Without chemotherapy, doctors expected Saikewicz to live for several weeks or perhaps several months, during which he would not experience severe pain or suffering. In not ordering treatment, a lower court considered "the quality of life available to him [Saikewicz] even if the treatment does bring about remission."

The Supreme Judicial Court of Massachusetts rejected the lower court's judgment that the value of life could be equated with one measure of the quality of life, namely, Saikewicz's lower quality of life because of mental retardation. Instead, the Supreme Judicial Court interpreted "the vague, and perhaps ill-chosen, term 'quality of life' … as a reference to the continuing state of pain and disorientation precipitated by the chemotherapy treatment."[58] It balanced prospective benefit against pain and suffering to reach the conclusion that the patient's interests supported a decision not to provide chemotherapy.

From a moral standpoint, we agree with the court's conclusion in this legal opinion, but the concept of "quality of life" needs further analysis. Some writers have argued that we should reject *moral* or otherwise *evaluative* judgments about quality of life and rely exclusively on *medical* indications for treatment decisions. For example, Paul Ramsey argues that for incompetent patients we need only determine which treatment is medically indicated to know which treatment is obligatory and which is optional. For imminently dying patients, responsibilities are not fixed by obligations to provide treatments that serve only to extend the dying process; they are fixed by obligations to provide appropriate care in dying. Ramsey predicts that unless we use these medical guidelines, we will gradually move toward a policy of active, involuntary euthanasia for unconscious or incompetent, nondying patients, based on arbitrary and inappropriate quality-of-life judgments.[59]

However, putatively objective medical factors, such as criteria used to determine medical indications for treatment, do not provide the objectivity that Ramsey seeks. These criteria undermine his fundamental distinction between the medical and the moral (or evaluative). It is impossible to determine what will benefit a patient without presupposing some quality-of-life standard and some conception of the life the patient will live after a medical intervention. Accurate medical diagnosis and prognosis are indispensable. But a judgment about whether to use life-prolonging measures rests unavoidably on the anticipated quality of life of the patient and cannot be reduced to vague and contestable standards of what is medically indicated.[60]

Ramsey maintains that a quality-of-life approach improperly shifts the focus from whether treatments benefit patients to whether patients' lives are beneficial to them—a shift that opens the door to active, involuntary euthanasia.[61] The underlying issue is whether we can state criteria of quality of life with sufficient precision and cogency to avoid such dangers. We think we often can, although the vagueness surrounding terms such as

dignity and *meaningful life* is a cause for concern, and cases in which seriously ill or disabled newborn infants have been "allowed to die" under questionable justifications also provide a reason for caution.

We should exclude several conditions of patients from consideration. For example, intellectual disability is irrelevant in determining whether treatment is in the patient's best interest. Furthermore, proxies should not confuse quality of life for the patient with the value of the patient's life for others. Instead, criteria focused on the incompetent patient's best interests should be decisive for a proxy, even if the patient's interests conflict with familial or societal interests in avoiding burdens or costs.

This position contrasts with that of the US President's Commission for the Study of Ethical Problems in Medicine and Biomedical and Behavioral Research, which recognized a broader conception of "best interests" that includes the welfare of the family: "The impact of a decision on an incapacitated patient's loved ones may be taken into account in determining someone's best interests, for most people do have an important interest in the well-being of their families or close associates."[62] Patients often do have an interest in their family's welfare, but it is a long step from this premise to a conclusion about whose interests should be overriding unless a competent patient explicitly so states. When the incompetent patient has never been competent or has never expressed his or her wishes while competent, it is improper to impute altruism or any other motive to that patient against his or her medical best interest.

Children with serious illnesses or disabilities. Endangered near-term fetuses and critically ill newborns or young children often pose difficult questions about medical treatment, particularly because of prognostic uncertainties concerning survival or quality of life. Prenatal obstetric management and neonatal intensive care can salvage the lives of many anomalous fetuses, premature babies, and newborns with physical conditions that would have been fatal a few decades ago. The reduction in infant mortality in the United States has been amazing, moving from an infant mortality rate of 25 deaths per 1,000 live births in 1960 to 5.74 deaths per 1,000 in 2014.[63] Celebrations of this success have been somewhat muted by serious concerns about the quality of life that some survivors face. Because the resultant quality of life is sometimes remarkably low, questions arise in some cases about whether aggressive obstetric management or intensive care will produce more harms and burdens than benefits for young patients.

As we argued at the end of Chapter 4 (pp. 141–42), the most appropriate standard in treatment decisions for never-competent patients, including critically ill newborns and young children, is that of best interests, as judged by the best estimate of what reasonable persons would consider the highest net benefit, in view of the probable benefits of different treatments balanced against their probable harms and burdens to the patients. Parents or other surrogates for these never-competent patients can legitimately use predictions about survival and about quality of life, evaluated according to the patients' interests, to determine whether treatments are obligatory, optional, or even, in extreme cases, wrong.

When a newborn or young child can be predicted to have such an extremely low quality of life following intensive care that the treatment can justifiably be judged to produce more harm than benefit, parents and the medical team are warranted in withholding or withdrawing treatment. Some conditions that arguably lead to a sufficiently poor quality of life to meet this standard include severe brain damage caused by birth asphyxia; Tay-Sachs disease, which involves increasing spasticity and dementia and usually results in death by age three or four; Lesch-Nyhan disease, which involves uncontrollable spasms, mental disability, compulsive self-mutilation, and early death; severe dystrophic epidermolysis bullosa, in which the child's skin inexorably peels off, resulting in excruciating pain and causing major infections that often kill the child in the first year of life, even with medical treatments. In some of these cases, particularly the last, it may even be wrong to treat because the anticipated short life with its abysmal quality could be reasonably assessed as caused by human intervention and as "intolerable."[64] A decision not to treat is also justifiable in severe cases of neural tube defects in which newborns lack all or most of the brain and will inevitably die. Premature babies at different gestational stages raise similar issues. One book in neonatal ethics maps these different stages by combining the best-interest standard with classificatory categories such as ours: obligatory to treat, optional, and obligatory not to treat.[65]

The best-interest standard, as a specification of principles of nonmaleficence and beneficence, focuses caregivers' attention on the interests of the newborn or young child, against other interests, including familial or societal interests. However, this approach does not preclude attention to these other interests in making ethical judgments. In the end, ethical judgments must take into account and balance the full range of important ethical considerations, including, for example, justice in the use of scarce resources as well as the overall best interests of the patient. Nevertheless, the best-interest standard serves as a *guidance* standard for the decisions of parents, who are the presumptive decision makers, and for the physicians and others who must provide information about possible options and their probable outcomes and must counsel parents.

The best-interest standard does not presuppose that there is always a single best plan for the newborn or young child. Where significant uncertainties are present about prognoses for survival or for quality of life—or where legitimate and reasonable differences exist in the values used to determine, weigh, and balance the patient's different interests, particularly as related to quality of life—different parents faced with the same situation may reasonably make different decisions. Parents usually have fairly wide latitude and discretion in making decisions about their children, for instance, regarding how to educate them, whether to allow them to engage in risky sports, and the like. The best-interest standard not only provides guidance in terms of the target (the child) and substance of the decision (the child's interests), but it also leaves room for parental discretion in many cases.

Some writers in bioethics argue that a "harm standard" is needed to supplant or supplement the best-interests standard for decision making about treatment for incapable patients such as newborns or infants.[66] In our judgment, this debate is misplaced, because the best-interests standard essentially incorporates the harm standard.[67] If an intervention is deemed to be in the patient's best interests, it is expected to provide a net benefit, considering the patient's interests in prolonged life, avoidance of pain and suffering, having a sufficient quality of life, and the like. This judgment rests on a probabilistic prediction of outcomes along with an evaluation of these outcomes through balancing or weighing different interests. If the intervention is not in the patient's best interests, providing it would often harm the patient and not merely fail to benefit him or her. An intervention against or contrary to the patient's overall interests sets back the patient's interests and thereby is harmful. When it is argued that avoidance of harm (including iatrogenic harm) is the most suitable guide to decisions on behalf of near-term fetuses and infants in neonatal care,[68] the assessment generally should be understood as avoiding a *net harm*. Most interventions inflict some harms, burdens, and the like on the patient but may still be in the patient's overall best interest.

The harm standard, as a subset of the best-interests standard, mainly provides a threshold for state intervention rather than a comprehensive guide for caregivers in their deliberations. This standard is and should be invoked when parents refuse treatments that are deemed by caregivers to be in an infant's best interests and caregivers seek a court order to override the parents' refusal. In these cases, the parental refusal to authorize a treatment in the infant's best interest is a setback to the patient's overall interests and therefore a net harm. Similar conclusions are in order for parental demands for treatments that are not in the patient's best interests. The harm standard does not supplant, replace, or supplement the best-interest standard. The best-interest standard, properly understood, incorporates the harm standard. (Later in this chapter we consider when it is justifiable to seek to disqualify parental or other surrogate decision makers.)

Debates about a newborn's or infant's best interest often surround parental *refusals* of treatments. The following case illustrates some complexities, ambiguities, uncertainties, and difficulties in the use of the best-interest standard.[69] Prenatal diagnosis detected fetal tricuspid atresia (TA), which is characterized by the absence of a tricuspid heart valve or the presence of an abnormal one; both conditions prevent blood flow from the right atrium to the right ventricle. In this particular case, the diagnosis of TA was made too late in the pregnancy for termination to be an option. The discussion centered on what to do after delivery. The cardiologist explained to the couple the nature of this condition—which can be relieved, but not cured, through immediate and long-term surgical and medical interventions—and the long-term prognosis. The cardiologist also discussed possible and probable morbidities and impacts on quality of life. The pregnant woman and her husband indicated that they wanted only end-of-life care after their baby's birth. Their decision was based in part on what they had learned from Internet searches, which showed that many parents refuse surgery for their infants under these circumstances.

A condition similar to TA is hypoplastic left heart syndrome (HLHS). In both situations the interventions are not curative. At the institution where this case occurred (as is true at most institutions in the United States), parents of newborns with HLHS may choose between surgery (the treatment also requires additional subsequent surgeries) and end-of-life care (data indicate that within the United States, parents are divided on this choice). Accordingly, the neonatologist argued that in order to treat equally situated patients equally, parents of infants with TA should also be able to choose between surgery and palliative care. The ethics debate was complicated because as many as 50% of infants with TA who do not receive early surgery live past the first year of life and some may even survive for several years, with the prospect of a long dying process together with significant distress and suffering. This risk of harm has led some to judge that it is justifiable to seek a court order for treatment against the parental refusal in this case.[70] An alternative approach allows the parents to refuse surgery, with full counseling about the possible outcomes, to be followed by a reevaluation of what to do if the infant survives for six months.[71]

Parental *requests* of treatments for infants also can be against their best interests, if the proposed treatments are (a) futile (as discussed earlier) or (b) have a low probability of benefit and a high probability of harm, including pain and suffering. The widely discussed case of Charlie Gard in the United Kingdom is an example of (a) and, according to a court opinion, (b) as well. As an eleven-month-old child, Charlie Gard had a rare condition, mitochondrial DNA depletion syndrome, which is uniformly fatal. He suffered epileptic seizures as well as discomfort related to intensive care, including ventilation, tube feeding, suctioning, and the like, all managed medically through treatments such as sedation and analgesia. It is unclear, and perhaps impossible to know, whether he experienced pain or pleasure or meaningful social interactions. Charlie Gard's parents wanted to try a highly experimental procedure—one never tried on his particular variant of the condition—in the United States, and they raised enough money to cover the costs. Nevertheless, the High Court in London ruled against them, holding that it was in their son's best interest for treatment to stop so he could die.[72]

In opposing this court decision, Julian Savulescu does not argue that this experimental treatment was in Charlie Gard's best interest, but only that it "is enough to say we don't know whether life will turn out to be in his interests and worth living."[73] Even though the odds of success were considered quite low, Savulescu does not see any acceptable grounds for denying him this chance at a decent life. By contrast, Dominic Wilkinson argues that the parents' request for treatment should not be allowed if no appropriately trained health professionals consider the experimental treatment worth pursuing. In this case, even the physician who was willing to provide the experimental treatment in the United States considered a benefit "unlikely."[74] Savulescu and Wilkinson agree, as do we, that there also might be grounds of distributive justice for denying this treatment option if public resources were required.

The fact that the court misapplied or overapplied the best-interest standard in this case should not be taken as an argument against the standard itself. Some critics construe this misapplication or overapplication as decisive evidence of the high susceptibility of the best-interest standard to value judgments and subjectivity.[75] Undeniably, this standard involves value judgments—notions of interests, best interests, harms, burdens, and the like often do—and subjectivity should be controlled or contained by imposing a requirement of reasonableness in those judgments. As vague and seemingly unwieldly as it sometimes appears, the best-interest standard remains the best standard for focusing parental and clinical deliberations about decisions to treat or to withhold or withdraw treatment from critically ill newborns and children. This standard also functions in some difficult and unresolvable conflicts to justify seeking a court order to override parental decisions that are sufficiently contrary to the newborn's or child's overall interests that they constitute a net harm.

Because the best-interest standard captures only one prima facie set of moral considerations connected to nonmaleficence and beneficence, other considerations such as distributive justice also enter into deliberations about the right course of action—a problem we consider in Chapter 7.

KILLING AND LETTING DIE

The distinction between killing and letting die (or allowing to die) is the most difficult and the most important of all of the distinctions that have been used to determine acceptable decisions about treatment and acceptable forms of professional conduct with seriously ill or injured patients. This distinction has long been invoked in public discourse, law, medicine, and moral philosophy to distinguish appropriate and inappropriate ways for death to occur. Killing has been widely viewed as morally wrong and letting die as morally acceptable. A large body of distinctions and rules about life-sustaining treatments derives from the killing-letting die distinction, which in turn draws on the act-omission and active-passive distinctions.[76] For instance, the killing-letting die distinction has affected distinctions between suicide (including assisted suicide) and forgoing treatment and between homicide and natural death.[77]

In considering whether this distinction is coherent, defensible, and useful for moral guidance, this section addresses three types of questions. (1) *Conceptual questions*: What conceptually is the difference between killing and letting die? (2) *Moral questions*: Is killing in itself morally wrong, whereas allowing to die is not in itself morally wrong? (3) *Combined conceptual and causal questions*: Is forgoing life-sustaining treatment sometimes a form of killing? If so, is it sometimes suicide and sometimes homicide?

Conceptual Questions about Killing and Letting Die

Can we define *killing* and *letting die* so that they are conceptually distinct and do not overlap? The following two cases suggest that we cannot: (1) A newborn with Down syndrome needed an operation to correct a tracheoesophageal fistula (a congenital deformity in which a connection exists between the trachea and the esophagus that allows food and milk to get into the lungs). The parents and physicians judged that survival was not in this infant's best interests and decided to let the infant die rather than undergo the operation. However, during a public outcry that erupted over this case, critics charged that the parents and physicians had killed the child by negligently allowing the child to die. (2) Dr. Gregory Messenger, a dermatologist, was charged with manslaughter after he unilaterally disconnected his fifteen-weeks premature (one-pound, 11-ounce) son's life-support system in a Lansing, Michigan, neonatal intensive care unit. Messenger thought he had merely acted compassionately in letting his son die after a neonatologist failed to fulfill a promise not to resuscitate the infant.[78]

Can we legitimately describe actions that involve intentionally not treating a patient as "allowing to die" or "letting die," rather than "killing"? Do at least some of these actions involve both killing and allowing to die? Is "allowing to die" a euphemism in some cases for "acceptable killing" or "acceptable ending of life"? These conceptual questions all have moral implications. Unfortunately, both ordinary discourse and legal concepts are vague and equivocal. In ordinary language, *killing* is a causal action that brings about death, whereas *letting die* is an intentional avoidance of causal intervention so that disease, system failure, or injury causes death. Killing extends to animal and plant life. Neither in ordinary language nor in law does the word *killing* entail a wrongful act or a crime, or even an intentional action. For example, we can say properly that in automobile accidents, one driver killed another even when no awareness, intent, or negligence was present.

Hence, conventional definitions are unsatisfactory for drawing a sharp distinction between killing and letting die. They allow many acts of letting die to count as killing, thereby defeating the point of the distinction. For example, under these definitions, health professionals kill patients when they intentionally let them die in circumstances in which they have a duty to keep the patients alive. It is unclear in literature on the subject how to distinguish killing from letting die so as to avoid even simple cases that satisfy the conditions of both killing and letting die. The meanings of "killing" and "letting die" are so vague and inherently contestable that attempts to refine their meanings likely will produce controversy without closure. We use these terms because they are prominent in mainstream literature, but we avoid a heavy reliance on them insofar as possible in the discussion below.

Connecting Judgments of Right and Wrong to Killing and Letting Die

"Letting die" is prima facie acceptable in medicine under one of two conditions: (1) a medical technology is *useless* in the strict sense of medical futility, as discussed earlier in this chapter, or (2) patients or their authorized surrogates have *validly refused* a medical technology. That is, letting a patient die is acceptable if and only if it satisfies the condition of futility or the condition of a valid refusal of treatment. If neither of these two conditions is satisfied, then letting a patient die constitutes killing (perhaps by negligence).

In medicine and health care, "killing" has traditionally been conceptually and morally linked to unacceptable acts. The conditions of medical practice make this connection understandable, but killing's absolute unacceptability is not assumed outside of specific settings such as traditional medical circles. The term *killing* does not necessarily entail a wrongful act or a crime, and the rule "Do not kill" is not an absolute rule. Standard justifications of killing, such as killing in self-defense, killing to rescue a person endangered by other persons' wrongful acts, and killing by misadventure (accidental, nonnegligent killing while engaged in a lawful act) prevent us from prejudging an action as wrong merely because it is a killing. Correctly applying the label "killing" or the label "letting die" to a set of events (outside of traditional assumptions in medicine) will therefore fail to determine whether an action is acceptable or unacceptable. There are both acceptable and unacceptable killings and both acceptable and unacceptable cases of allowing to die.[79]

It may be that killing is usually wrong and letting die only rarely wrong, but, if so, this conclusion is contingent on the features of particular cases. The general wrongness of killing and the general rightness of letting die are not surprising features of the moral world inasmuch as killings are rarely authorized by appropriate parties (excepting contexts such as warfare and capital punishment) and cases of letting die generally are validly authorized. Be that as it may, the *frequency* with which one kind of act is justified, in contrast to the other kind of act, cannot determine whether either kind of act is legally or morally justified in particular cases. Forgoing treatment to allow patients to die can be both as intentional and as immoral as actions that in some more direct manner take their lives, and both can be forms of killing.

In short, the labels "killing" and "letting die," even when correctly applied, do not determine that one form of action is better or worse, or more or less justified, than the other. Some particular instance of killing, such as a brutal murder, may be worse than some particular instance of allowing to die, such as forgoing treatment for a PVS patient; but some particular instance of letting die, such as not resuscitating a patient whom physicians could potentially save, also may be worse than some particular instance of killing, such as mercy killing at the patient's request. Nothing about either killing or allowing to die entails judgments about actual wrongness or rightness. Rightness and wrongness depend on the merit of the justification underlying the action, not on whether it is an instance of killing or of letting die. Neither killing nor letting die is per se wrongful, which distinguishes them from murder, which is per se wrongful.

Accordingly, judging whether an act of either killing or letting die is justified or unjustified requires that we know something else about the act besides these characteristics. We need to know about the circumstances, the actor's motive (e.g., whether it is benevolent or malicious), the patient's preferences, and the act's consequences. These additional factors will allow us to place the act on a moral map and make an informed normative judgment about whether it is justifiable.

Forgoing Life-Sustaining Treatment: Killing or Allowing to Die?

Many writers in medicine, law, and ethics have construed a physician's intentional forgoing of a medical technology as letting die if and only if an underlying disease or injury causes death. When physicians withhold or withdraw medical technology, according to this interpretation, a natural death occurs, because natural conditions do what they would have done if the physicians had never initiated the technology. By contrast, killings occur when acts of persons rather than natural conditions cause death. From this perspective, one acts nonmaleficently in allowing to die and maleficently in killing (whatever one's motives may be).

Although this view is influential in law and medicine, it is flawed. To attain a satisfactory account, we must add that the forgoing of the medical technology is *validly authorized* and *for this reason justified*. If the physician's forgoing of technology were unjustified and a person died from "natural" causes of injury or disease, the result

would be unjustified killing, not justified allowing to die. The validity of the authorization—not some independent assessment of the causation of death—*determines the moral acceptability of the action*. For example, withdrawing treatment from a competent patient is not morally justifiable unless the patient has made an informed decision authorizing this withdrawal. If a physician removes a respirator from a competent patient who needs it and wants to continue its use, the action is wrong, even though the physician has only removed artificial life support and let nature take its course. The lack of authorization by the patient is the relevant consideration in assessing the act as unacceptable, not the distinction between letting die and killing.

Even from a legal perspective, we can provide a better causal account than "the preexisting disease caused the death." The better account is that legal liability should not be imposed on physicians and surrogates unless they have an obligation to provide or continue the treatment. If no obligation to treat exists, then questions of causation and liability do not arise. If the categories of obligatory and optional are primary, we have a reason for avoiding discussions about killing and letting die altogether and for focusing instead on health care professionals' obligations and problems of moral and legal responsibility.

In conclusion, the distinction between killing and letting die suffers from vagueness and moral confusion. Specifically, the language of killing and its use in much of the literature of biomedical ethics is sufficiently confusing—causally, legally, and morally—that it provides little, if any, help in discussions of assistance in dying. In the next section we further support this conclusion.

Intentionally Arranged Deaths: When, If Ever, Are They Justified?

We now address a set of moral questions about the causation of death that are largely free of the language of "killing." The general question is, "Under which conditions, if any, is it permissible for a patient and a health professional to arrange for the health professional's assistance in intentionally ending the patient's life?"

Withholding or withdrawing treatment will hasten death only for individuals who could be or are being sustained by a technology. Many other individuals, including some patients with cancer, face a protracted period of dying when respirators and other life-preserving technology are not being utilized. Great improvements in and extensions of palliative care can adequately address the needs of many, perhaps most, of these patients.[80] However, for many others, palliative care and the refusal of particular treatments do not adequately address their concerns. During their prolonged period of dying, they may endure a loss of functional capacity, unremitting pain and suffering, an inability to experience the simplest of pleasures, and long hours aware of the hopelessness of their condition. Some patients find this prospect, or its actuality, unbearable and desire a painless means to hasten their deaths.

In addition to withholding or withdrawing treatments or technologies, and prescribing medications that may relieve pain and suffering while indirectly hastening death (see our discussion at pp. 167 and 170 of the rule of double effect), physicians sometimes use what is viewed as a more active means to bring about a patient's death. Some argue that the use of an active means in medicine to bring about death always constitutes an inappropriate killing, but there are problems in the idea that we can determine appropriate and inappropriate conduct by considering whether an active means was involved.

An example is the Oregon Death with Dignity Act (ODWDA),[81] where the distinction between "letting die" and "killing" is not used and, in any event, would not be helpful in addressing particular cases under this act. Physicians who act under the terms of ODWDA do not "kill" when acting as permitted under the law; rather, they write prescriptions for a lethal medication at a patient's request. The patient must make a conscious decision whether to use the drug. As many as one-third of the patients who receive a written prescription never ingest the lethal drug. For those who take the drug, the physician's writing of the prescription is a necessary step in the process that leads to some patients' deaths, but it is not the determinative or even the final step, and so is not the cause of a patient's death. Under any reasonable interpretation of the term, the Oregon physician does not

"kill" the patient, nor does a physician "let the patient die." Here the terms *letting die* and *killing* do not illuminate or help evaluate what happens when a physician helps a person escape the ravages of a fatal illness.

Some literature in bioethics treats issues about active physician assistance under the umbrella of the legal protection of a "right to die."[82] Underlying the legal issues is a powerful struggle in law, medicine, and ethics over the nature, scope, and foundations of the right to choose the manner of one's death. Below we discuss legalization, public policy, and institutional policy, but we are primarily interested in whether acts of assistance by health professionals are *morally justified*. We begin with an important distinction between acts and policies. From there we work back to some foundational moral issues.

Acts, Practices, and Slippery Slopes

Justifying an act is distinct from justifying a practice or a policy that permits or even legitimates the act's performance. A rule of practice or a public policy or a law that prohibits various forms of assistance in dying in medicine may be justified even if it excludes some acts of causing a person's death that in themselves, as acts, are *morally* justified. For example, sufficient reasons may justify a law in a particular jurisdiction that prohibits physicians from prescribing a lethal drug. However, in a particular case in that jurisdiction, it could be ethically justifiable to provide the drug to a patient who suffers from terrible pain, who will probably die within a few weeks, and who requests a merciful assisted death. In short, a valid and ethically justified law might forbid an action that is morally justified in some individual cases.

A much-discussed problem is that a practice or policy that allows physicians to intervene to cause deaths or to prescribe lethal drugs runs risks of abuse and might cause more harm than benefit. The argument is not that serious abuses will occur immediately, but that they will grow incrementally over time. Society could start by severely restricting the number of patients who qualify for assistance in dying, but later loosen these restrictions so that cases of unjustified killing begin to occur. Unscrupulous persons would learn how to abuse the system, just as they do now with methods of tax evasion on the margins of the system of legitimate tax avoidance. In short, the argument is that the slope of the trail toward the unjustified taking of life could be so slippery and precipitous that we ought never to embark on it.

Many dismiss such slippery-slope, or wedge, arguments because of a lack of empirical evidence to support the claims involved, as well as because of their heavily metaphorical character ("the thin edge of the wedge," "the first step on the slippery slope," "the foot in the door," and "the camel's nose under the tent"). However, some slippery-slope arguments should be taken seriously in certain contexts.[83] They force us to think about whether unacceptable harms or wrongs may result from attractive, and apparently innocent, first steps. If society removes certain restraints against interventions that cause death, various psychological and social forces could make it more difficult to maintain the relevant distinctions in practice.

Opponents of the legalization of physician-assisted dying have often maintained that the practice inevitably would be expanded to include euthanasia, that the quality of palliative care for all patients would deteriorate, that patients would be manipulated or coerced into requesting assistance in hastening death, that patients with impaired judgment would be allowed to request such assistance, and that members of possibly vulnerable groups (people with disabilities, the economically disadvantaged, the elderly, immigrants, members of racial and ethnic minorities, etc.) would be adversely affected in disproportionate numbers. These slippery-slope claims are credible in light of the effects of social discrimination based on disability, cost-cutting measures in the funding of health care, and the growing number of elderly persons with medical problems that require larger and larger proportions of a family's or the public's financial resources. If rules allowing physician-assisted dying became public policy, the risk would increase that persons in these populations will be neglected or otherwise abused. For example, the risk would increase that some families and health professionals would abandon treatments for disabled newborns and adults with severe brain damage to avoid social and familial burdens. If decision makers reach judgments that some newborns and adults have overly burdensome conditions or lives with no value, the same logic can be extended to populations of feeble, debilitated, and seriously ill patients who are financial and emotional burdens on families and society.

These fears are understandable. Rules in a moral code against passively or actively causing the death of another person are not isolated fragments. They are threads in a fabric of rules that uphold respect for human life. The more threads we remove, the weaker the fabric might become. If we focus on the modification of attitudes and beliefs, not merely on rules, shifts in public policy may also erode the general attitude of respect for life. Prohibitions are often both instrumentally and symbolically important, and their removal could weaken critical attitudes, practices, and restraints.

Rules against bringing about another's death also provide a basis of trust between patients and health care professionals. We expect health care professionals to protect and promote our welfare under all circumstances. We may risk a loss of public trust if physicians become agents of intentionally causing death in addition to being healers and caregivers. On the other side, however, we may also risk a loss of trust if patients and families believe that physicians abandon them in their suffering because the physicians lack the courage to offer the assistance needed in the darkest hours of their lives.[84]

Slippery-slope arguments ultimately depend on speculative predictions of a progressive erosion of moral restraints. If dire consequences will probably flow from the legalization of physician-assisted dying in a jurisdiction, then these arguments are cogent and it is justifiable to prohibit such practices in that jurisdiction. But how good is the evidence that dire consequences will occur? Does the evidence indicate that we cannot maintain firm distinctions in public policies between, for example, patient-requested death and involuntary euthanasia?[85]

Scant evidence supports the many answers that have been given to these questions. Those of us, including the authors of this book, who take seriously the cautions presented in some versions of the slippery-slope argument should admit that it requires a premise on the order of a precautionary principle, such as "better safe than sorry." (See our discussion of a precautionary approach and process in Chapter 6.) The likelihood of the projected moral erosion is not something we presently can assess by appeal to good evidence. Arguments on every side are speculative and analogical, and different assessors of the same evidence reach different conclusions. Intractable controversy likely will persist over what counts as good and sufficient evidence. How Oregon's procedural safeguards work, or fail to work, will continue to be carefully watched. That state's experience has influenced subsequent steps taken in other states and countries. Failure of the ODWDA would be a major setback for proponents of the right to die by use of prescribed drugs.

However, two decades after the enactment of the Oregon law, none of the abuses some had predicted materialized in Oregon.[86] The Oregon statute's restrictions have been neither loosened nor broadened. There is no evidence that any patient has died other than in accordance with his or her own wishes. While the number of patients receiving prescriptions under the statute has increased significantly (from 24 in 1998 to 88 in 2008 to 218 in 2017), the law has not been used primarily by individuals who might be thought vulnerable to intimidation or abuse. Those choosing assisted death have had, on average, a higher level of education and better medical coverage than terminally ill Oregonians who did not seek assistance in dying. Women, people with disabilities, and members of disadvantaged racial minorities have not sought assistance in dying in disproportionate numbers. The overwhelming number of persons requesting assistance in dying are Caucasian, and the gender of the requesters reflects the general population. Meanwhile, reports indicate that the quality of palliative care has improved in Oregon. In 2017 approximately 20% of the 218 patients receiving a prescription for a lethal medication decided not to use the prescribed drug (at least during 2017); data were not confirmed about use or nonuse for an additional 20% (at the time of the annual report).[87]

Oregon's experiment in physician-assisted death is instructive and reassuring in many respects, but questions inevitably arise about its generalizability as a model for the whole of the United States and for other countries, just as they arise about experiments with assisted dying in countries such as the Netherlands, Belgium, Canada, and Switzerland.[88]

Valid Requests for Aid-in-Dying

We now go to the central question of whether some acts of assisting another in dying are morally justified and others unjustified. The frontier of expanded rights to control one's death shifted, roughly at the point of the transition from the twentieth to the twenty-first century, from *refusal* of treatment to *requests* for aid-in-dying.[89] Assuming that the principles of respect for autonomy and nonmaleficence justify forgoing treatment, the same justification, coupled with the principle of beneficence, might be extended to physicians prescribing barbiturates or providing other forms of help requested by seriously ill patients. This strategy relies on the premise that professional ethics and legal rules should avoid the apparent inconsistency between (1) the strong rights of autonomous choice that allow persons in grim circumstances to refuse treatment in order to bring about their deaths and (2) the denial of a similar autonomy right for persons under equally grim circumstances to arrange for death by mutual agreement with a physician. The argument for reform is compelling when a condition overwhelmingly burdens a patient, pain management fails to adequately comfort the patient, and only a physician can and is willing to bring relief. At present, medicine and law in most jurisdictions in the United States are in the awkward position of having to say to such patients, "If you were on life-sustaining treatment, you would have a right to withdraw the treatment and then we could let you die. But since you are not, we can only allow you to refuse nutrition and hydration or give you palliative care until you die a natural death, however painful, undignified, and costly."[90]

The two types of autonomous action—refusal of treatment and request for aid-in-dying—are not perfectly analogous. A health professional is firmly obligated to honor an autonomous refusal of a life-prolonging technology, but he or she is not obligated under ordinary circumstances to honor an autonomous request for aid-in-dying. The key issue is not whether physicians are morally *obligated* to lend assistance in dying, but whether valid requests render it morally *permissible* for a physician (or possibly some person other than a physician) to lend aid-in-dying. Refusals in medical settings generally have a moral force not found in requests, but requests do not lack all power to confer on another person a right to perform the requested act.

A physician's precise responsibilities to a patient may depend on the nature of the request made as well as on the preestablished patient-physician relationship. In some cases of physician compliance with requests, the patient and the physician pursue the patient's best interest under an agreement that the physician will not abandon the patient and will undertake to serve what they jointly determine to be the patient's best interests. In some cases, patients in a close relationship with a physician both refuse a medical technology and request a hastened death to lessen pain or suffering. Refusal and request may be two parts of a single inclusive plan. If the physician accepts the plan, some form of assistance grows out of the preestablished relationship. From this perspective, a valid request for aid-in-dying frees a responder of moral culpability for the death, just as a valid refusal precludes culpability.

These arguments suggest that causing a person's death is morally wrong, when it is wrong, because an unauthorized intervention thwarts or sets back a person's interests. It is an unjustified act when it deprives the person who dies of opportunities and goods.[91] However, if a person freely authorizes his or her death by making an autonomous judgment that ending life because of a need to diminish pain and suffering, an inability to engage in activities making life enjoyable, a reduced autonomy or dignity, a loss of control of bodily functions, or being a burden on one's family constitutes a personal benefit rather than a setback to interests, then active aid-in-dying at the person's request involves neither harming nor wronging.[92] Aiding an autonomous person at his or her request for assistance in dying is, from this perspective, a way of showing respect for the person's autonomous choices. Similarly, denying the person access to individuals who are willing and qualified to comply with the request can show a fundamental disrespect for the person's autonomous choice.

Unjustified Physician Assistance in Dying

The fact that the autonomous requests of patients for aid-in-dying should be respected in some circumstances does not entail that *all* cases of physician-assisted death at the patient's request are justifiable. Jack Kevorkian's widely reported practices provide an important historical example of the kind of *unjustified* physician assistance that society should discourage and even prohibit. In his first case of assisting in suicide, Janet Adkins, an Oregon grandmother with Alzheimer's disease, had reached a decision that she wanted to take her life rather than lose

her cognitive capacities, which she was convinced were slowly deteriorating. After Adkins read in news reports that Kevorkian had invented a "death machine," she communicated with him by phone and then flew from Oregon to Michigan to meet with him. Following brief discussions, she and Kevorkian drove to a park in northern Oakland County. He inserted a tube in her arm and started saline flow. His machine was constructed so that Adkins could then press a button to inject other drugs, culminating in potassium chloride, which would physically cause her death. She then pressed the button.[93]

This case raises several concerns. Janet Adkins was in the fairly early stages of Alzheimer's and was not yet debilitated. At fifty-four years of age, she was still capable of enjoying a full schedule of activities with her husband and playing tennis with her son, and she might have been able to live a meaningful life for several more years. A slight possibility existed that the Alzheimer's diagnosis was incorrect, and she might have been more psychologically depressed than Kevorkian appreciated. She had limited contact with him before they collaborated in her death, and he did not administer examinations to confirm either her diagnosis or her level of competence to commit suicide. Moreover, he lacked the professional expertise to evaluate her medically or psychologically. The glare of media attention also raises the question whether Kevorkian acted imprudently to generate publicity for his social goals and for his forthcoming book.

Lawyers, physicians, and writers in bioethics have almost universally condemned Kevorkian's actions. The case raises all the fears present in the arguments mentioned previously about physician-assisted dying: lack of social control, inadequate medical knowledge, unconfirmed medical diagnoses and prognoses, no serious and qualified assessment of the patient's mental and emotional state, absence of accountability, and unverifiable circumstances of a patient's death. Although Kevorkian's approach to assisted suicide was deplorable, some of his "patients" raised distressing questions about the lack of a support system in health care for handling their problems. Having thought for over a year about her future, Janet Adkins decided that the suffering of continued existence exceeded its benefits. Her family supported her decision. She faced a bleak future from the perspective of a person who had lived an unusually vigorous life, both physically and mentally. She believed that her brain would slowly deteriorate, with progressive and devastating cognitive loss and confusion, fading memory, immense frustration, and loss of all capacity to take care of herself. She also believed that the full burden of responsibility for her care would fall on her family. From her perspective, Kevorkian's offer was preferable to what other physicians had offered, which was a flat refusal to help her die as she wished.

Justified Physician Assistance in Dying

Kevorkian's strategy is an example of *unjustified* assisted suicide. By contrast, consider the actions of physician Timothy Quill in prescribing the barbiturates desired by a forty-five-year-old patient who had refused a risky, painful, and often unsuccessful treatment for leukemia. She had been his patient for many years and she and members of her family had, as a group, come to this decision with his counsel. She was competent and had already discussed and rejected all available alternatives for the relief of suffering. This case satisfied the general conditions that are sufficient for justified physician assistance in ending life. These conditions, we propose, include

1. 1. A voluntary request by a competent patient
2. 2. An ongoing patient-physician relationship
3. 3. Mutual and informed decision making by patient and physician
4. 4. A supportive yet critical and probing environment of decision making
5. 5. A patient's considered rejection of alternatives
6. 6. Structured consultation with other parties in medicine
7. 7. A patient's expression of a durable preference for death
8. 8. Unacceptable suffering by the patient
9. 9. Use of a means that is as painless and comfortable as possible

Quill's actions satisfied all of these conditions, but critics found his involvement as a physician unsettling and unjustified. Several critics invoked slippery-slope arguments, because acts like Quill's, if legalized, could potentially affect many patients, especially the elderly. Others were troubled by the fact that Quill apparently

violated a New York state law against assisted suicide. Furthermore, to reduce the risks of criminal liability, Quill apparently lied to the medical examiner by informing him that a hospice patient had died of acute leukemia.[94]

Despite these problems, we do not criticize Quill's basic intentions in responding to the patient, the patient's decision, or their relationship. Suffering and loss of cognitive capacity can ravage and dehumanize patients so severely that death is in their best interests. In these tragic situations—or in anticipation of them, as in this case—physicians such as Quill do not act wrongly in assisting competent patients, at their request, to bring about their deaths. Public policy issues regarding how to avoid abuses and discourage and prevent unjustified acts should be a central part of our discussion of forms of appropriate physician assistance, but these problems do not finally determine the moral justifiability of the physician's act of assisting in the patient's death when caring for the patient.

Such caring physician assistance in hastening death is best viewed as part of a continuum of medical care. A physician who encounters a sick patient should initially seek, if possible, to rid the patient's body of its ills. Restoration of health is a morally mandatory goal if a reasonable prospect of success exists and the patient supports the means necessary to this end. However, to confine the practice of medicine to measures designed to cure diseases or heal injuries is an unduly narrow way of thinking about what the physician has to offer the patient. When, in the patient's assessment, the burdens of continued attempts to cure outweigh their probable benefits, a physician should be able to redirect the course of treatment so that its primary focus is the relief of pain and suffering. For many patients, palliative care with aggressive use of analgesics will prove sufficient to accomplish this goal. For other patients, relief of intolerable suffering will come only with death, which some will seek to hasten.

A favorable response by a physician to a request for assistance in facilitating death by *hastening* it through prescribing lethal medication is not relevantly different from a favorable response to requests for assistance in facilitating death by *easing* it through removal of life-prolonging technology or use of coma-inducing medications. The two acts of physician assistance are morally equivalent as long as no other morally relevant differences are present in the cases. That is, if in both cases the diseases are relevantly similar, the requests by the patient are relevantly similar, and the desperateness of the patients' circumstance is relevantly similar, responding to a request to provide the means to hasten death is morally equivalent to responding to a request to ease death by withdrawing treatment, sedating to coma, and the like.

With due caution, we should be able to devise social policies and laws that maintain a bright line between justified and unjustified physician assistance in dying. Principles of respect for autonomy and beneficence and virtues of care and compassion all offer strong reasons for recognizing the legitimacy of physician-assisted death. Major opposition stems from interpretations of the principle of nonmaleficence and its specifications in various distinctions and rules. We have argued that the most critical distinctions and rules often break down on closer examination. In arguing for changes in laws and policies to allow physician-assisted dying in certain contexts, we do not maintain that these changes will handle all important issues in the care of dying and seriously ill patients. The changes we recommend mainly address last-resort situations, which can often be avoided by better social policies and practices, including improved palliative care, which we also strongly recommend.

In presenting a case involving the disconnection of a ventilator maintaining the life of a patient with amyotrophic lateral sclerosis (ALS, or Lou Gehrig's disease) at an international conference on "Ethical Issues in Disability and Rehabilitation," some clinicians framed it as an "end-of-life case," in which the "patient" decided to discontinue the ventilator. They were surprised when the audience, many of whom had disabilities and had themselves experienced long-term ventilator use, disputed this classification and argued instead that this was a "disability" case in which the clinicians should have provided better care, fuller information, and more options to the "consumer," particularly to help him overcome his felt isolation after the recent death of his spouse: "What to the clinicians was a textbook case of 'end-of-life' decision making was, for their audience, a story in which a life was ended as a result of failures of information and assistance by the presenters themselves."[95]

Few doubt that we need further improvements in supporting people who suffer from serious medical problems. Control of pain and suffering is a moral imperative. However, significant progress in control of pain and suffering will not obviate last-resort situations in which individuals reasonably seek to control their dying in ways that have often been denied to patients.

PROTECTING INCOMPETENT PATIENTS FROM HARM

Laws that authorize physician-assistance in dying should apply only to competent persons who are able to make autonomous choices. There is vigorous debate about whether comparable laws should be extended to previously competent persons who have provided a clear and relevant advance directive. Apart from physician aid-in-dying, we have noted other possible decisions that may apply to incompetent persons, including newborns and children. In Chapter 4 (pp. 139–41), we examined *standards* of surrogate decision making for incompetent patients. We now consider *who* should decide for the incompetent patient. Determining the best system for protecting patients from harm is the central problem.[96] In the absence of advance directives executed by previously competent individuals, we think first of families as the proper decision makers because they usually have the deepest interest in protecting their incompetent members. However, we also need a system that will shield incompetent individuals from family members who care little or are caught in conflicts of interest, while at the same time protecting residents of nursing homes, psychiatric hospitals, and facilities for the disabled and mentally handicapped, many of whom rarely, if ever, see a family member. The appropriate roles of families, courts, guardians, conservators, hospital committees, and health professionals all merit consideration.

Advance Directives

In an increasingly popular procedure rooted as much in respect for autonomy as in obligations of nonmaleficence, a person, while competent, either writes a directive for health care professionals or selects a surrogate to make decisions about life-sustaining treatments during periods of incompetence.[97] Two types of *advance directive* aim at governing future decisions: (1) *living wills*, which are substantive or instructional directives regarding medical procedures in specific circumstances, and (2) *durable power of attorney* (DPA) for health care, which is a legal document that allows persons to assign a specific agent (a proxy or surrogate) to make their health care decisions when they have lost capacity. The power is "durable" because, unlike the usual power of attorney, it continues in effect when the signer becomes incompetent.

However, these documents generate practical and moral problems.[98] First, relatively few persons compose them, and many who do fail to leave sufficiently explicit instructions. Second, a designated decision maker might be unavailable when needed, might be incompetent to make good decisions for the patient, or might have a conflict of interest such as a prospective inheritance or a better position in a family-owned business. Third, some patients who change their preferences about treatment fail to change their directives, and a few legally incompetent patients protest a surrogate's decision. Fourth, laws in some legal jurisdictions severely restrict the use of advance directives. For example, advance directives may have legal effect if and only if the patient is terminally ill and death is imminent. However, difficult decisions often must be made when the patient is not imminently dying or does not have a medical condition appropriately described as a terminal illness. Fifth, living wills provide no basis for health professionals to overturn a patient's instructions; yet prior decisions by the patient could turn out not to be in the patient's best medical interest. Patients while competent often could not have reasonably anticipated the precise circumstances they actually encountered when they became incompetent. Surrogate decision makers also sometimes make decisions with which physicians sharply disagree, in some cases asking the physician to act against his or her conscience or against medical practice standards.

Despite these problems, the advance directive is a valid way for competent persons to exercise their autonomy, and implementing the procedures for informed consent discussed in Chapter 4 can overcome many of the practical problems. As in informed consent situations, we should distinguish the *process* from the *product* (here, the advance directive). Efforts are under way to enhance the entire process of advance care planning, for instance, through in-depth dialogue, improved communication, values histories, and the use of a variety of

scenarios and decision aids.[99] In contrast to earlier studies that found little if any impact of advance directives on subsequent decisions and care,[100] later research indicates that elderly patients who lose their capacity to make decisions but who have advance directives tend to receive care that is strongly aligned with their previously stated preferences. However, some studies indicate that advance directives have not significantly enhanced physician-patient communication and decision making about subjects such as resuscitation.[101]

Surrogate Decision Making without Advance Directives

When an incompetent patient lacks an advance directive, who should make which decisions, and with whom should the decision maker consult?

Qualifications of surrogate decision makers. We propose the following list of qualifications for decision makers for incompetent patients (including newborns):

1. 1. Ability to make reasoned judgments (competence)
2. 2. Adequate knowledge and information
3. 3. Emotional stability
4. 4. A commitment to the incompetent patient's interests, free of conflicts of interest and free of controlling influence by those who might not act in the patient's best interests

The first three conditions follow from our discussions of informed consent and competence in Chapter 4. The only potentially controversial condition is the fourth. Here we endorse a criterion of *partiality*—acting as an advocate in the incompetent patient's best interests—rather than *impartiality*, which requires neutrality when considering the interests of the various affected parties. Impartial consideration of the interests of all parties is not appropriate to the role of being an advocate for the patient.

Four classes of decision makers have been proposed and used in cases of withholding and terminating treatment for incompetent patients: families, physicians and other health care professionals, institutional committees, and courts. If a court-appointed guardian exists, that person will act as the primary responsible party. The following analysis is meant to provide a defensible structure of decision-making authority that places the caring family as the presumptive authority when the patient cannot make the decision and has not previously designated a decision maker.

The role of the family. Wide agreement exists that the patient's closest family member is the first choice as a surrogate. Many patients strongly prefer family members to interact with physicians as the decision-making authorities about their medical fate.[102] The family's role should be presumptively primary because of its presumed identification with the patient's interests, depth of concern about the patient, and intimate knowledge of his or her wishes, as well as its traditional position in society.

Unfortunately, the term *family* is imprecise, especially if it includes the extended family. The reasons that support assigning presumptive priority to the patient's closest family member(s) also support assigning relative priority to other family members. However, even the patient's closest family members sometimes make unacceptable decisions, and the authority of the family is not final or ultimate.[103] The closest family member can have a conflict of interest, can be poorly informed, or can be too distant personally and even estranged from the patient.[104]

Consider an illustrative case: Mr. Lazarus was a fifty-seven-year-old male patient brought into the hospital after suffering a heart attack while playing touch football. He lapsed into a coma and became ventilator-dependent. After twenty-four hours his wife requested that the ventilator be withdrawn and dialysis stopped to allow him to die. The attending physician was uncomfortable with this request because he thought that Mr. Lazarus had a good chance of full recovery. Mrs. Lazarus insisted that treatment be withdrawn, and she had a DPA for health care that designated her the surrogate. She became angry when the health care team expressed its reluctance to withdraw care, and she threatened to sue the hospital if her decision was not honored. An ethics consult was

called because the attending and staff remained unwilling to carry out her wishes. The ethics consultant read the DPA and discovered that Mr. Lazarus had designated his wife as surrogate only if he was deemed to be in a PVS. Furthermore, Mr. Lazarus had stipulated on the DPA that if he was not in a PVS, he wanted "everything done." He awoke after three days and immediately revoked his DPA when told of his wife's demand.[105]

Health care professionals should seek to disqualify any decision makers who are significantly incompetent or ignorant, are acting in bad faith, or have a conflict of interest. Serious conflicts of interest in the family may be more common than either physicians or the courts have generally appreciated.[106] Health care professionals also should be alert to and help address the burdens of decision making for familial and other surrogates. According to one review of the relevant research, at least one-third of the surrogates involved in decision making about treatment for incapacitated adults experienced emotional burdens, such as stress, guilt, and doubt about whether they had made the best decisions in the circumstances. However, when surrogates were confident that the treatment decision accorded with the patient's own preferences, their emotional burden was reduced.[107]

The role of health care professionals. Physicians and other health care professionals can help family members become more adequate decision makers and can safeguard the patient's interests and preferences, where known, by monitoring the quality of surrogate decision making. Physicians sometimes best serve both the family and the patient by helping surrogates see that rapid functional decline has set in and the time has come to shift from life-prolonging measures to palliative care centered on increasing comfort and reducing the burdens of treatments.[108] Such a reorientation can be wrenchingly difficult and emotionally challenging for physicians, nurses, and family members.

In the comparatively rare situation in which physicians contest a surrogate's decision and disagreements persist, an independent source of review, such as a hospital ethics committee or the judicial system, is advisable. In the event that a surrogate, a member of the health care team, or an independent reviewer asks a caregiver to perform an act the caregiver regards as contraindicated, futile, or unconscionable, the caregiver is not obligated to perform the act but may still be obligated to help the surrogate or patient make other arrangements for care.

Institutional ethics committees. Surrogate decision makers sometimes refuse treatments that would serve the interests of those they should protect, and physicians sometimes too readily acquiesce in their preferences. In other cases, surrogates need advice or help in reaching difficult decisions. The involved parties then may need a mechanism or procedure to help make a decision or to break a private circle of refusal and acquiescence. A similar need exists for assistance in decisions regarding residents of nursing homes and hospices, psychiatric hospitals, and residential facilities in which families often play only a small role, if any.

Institutional ethics committees can help in these situations, though they differ widely in their composition, function, and responsibilities. Many committees create or recommend explicit policies to govern actions such as withholding and withdrawing treatment, and many serve educational functions in hospitals or other institutions. Controversy centers on various additional functions, such as whether committees should make, facilitate, or monitor decisions about patients in particular cases. The decisions of committees on occasion need to be reviewed or criticized, perhaps by an auditor or impartial party.

Nonetheless, the benefits of good committee review generally outweigh its risks, and these committees have a robust role to play in circumstances in which physicians acquiesce too readily to parental, familial, or guardian choices that prove contrary to a patient's best interests.

The judicial system. Courts are sometimes unduly intrusive as final decision makers, but in many cases they represent the last and perhaps the fairest recourse. When good reasons exist to appoint guardians or to disqualify familial decision makers or health care professionals to protect an incompetent patient's interests, the courts may legitimately be involved. The courts also sometimes need to intervene in nontreatment decisions for incompetent patients in mental institutions, nursing homes, and the like. If no family members are available or willing to be involved, and if the patient is confined to a state mental institution or a nursing home, it may be appropriate to establish safeguards beyond the health care team and the institutional ethics committee.[109]

WHOSE RISKS AND WHOSE BENEFITS? PROBLEMS OF UNDERPROTECTION AND OVERPROTECTION IN RESEARCH

We have thus far concentrated on harm in clinical care. We now turn to ethical issues of harm in research.

Historical Problems of Underprotection

Historically, the risks of harm to human subjects in medical research have often been placed heavily on the economically disadvantaged, the very sick, and the vulnerable, owing to their ready availability. The unjustified overutilization of members of these populations has been a matter of deep moral concern in biomedical ethics. Even though there is general agreement that we need a system of research ethics with sufficient internal controls to protect subjects from exploitation, disagreement surrounds questions about the conditions under which protections are needed and how best to ensure those protections. In the last three decades of the twentieth century the predominant concern was that we were underprotecting human subjects, especially vulnerable groups such as children, the mentally handicapped, and the institutionalized. The harms caused by the underprotection of research subjects have been well documented and carefully examined in the biomedical ethics literature, and they have often been addressed in public policy and regulation as well.[110] However, the harms caused by the overprotection of subjects have received far less attention, even though they can create serious delays in the progress of research, thereby causing harm to those who do not receive the medical benefits of the research in a timely fashion. We emphasize this problem in the following subsection.

Recent Problems of Overprotection

An eye-opening case of such problems starts with an allegation of inappropriate human-subjects research on catheter-related bloodstream infections, which can cause thousands of deaths each year in intensive care units (ICUs).[111] Dr. Peter Pronovost, then at The Johns Hopkins University, was working with 103 ICUs in 67 Michigan hospitals to implement and evaluate what Johns Hopkins and other ICUs had established to be a successful infection-control measure. The work was halted by federal regulators in the Office for Human Research Protections (OHRP) after receiving a complaint that Pronovost and the hospitals were using patients in human-subjects research without their informed consent.

Pronovost's activities were part of a study to improve medical care sponsored by the Michigan Hospital Association. The aim was to control infections in ICUs by strictly implementing preventive procedures that had already been recommended by the Centers for Disease Control and Prevention, such as washing hands, using infection control precautions, and the like. The team studied the effect on infection rates of a careful implementation in practice of all the recommended procedures, following a checklist. They found that infection rates fall substantially when the checklist is scrupulously followed.

A published report of the study led to a complaint to the OHRP that the research violated US federal regulations. After investigating, the OHRP demanded that Johns Hopkins and the Michigan hospitals correct their "mistake" and undertake a full ethics review of the study. The Johns Hopkins institutional review board (IRB) had already examined the project and found that full IRB review and informed consent were *not* required in this case. This IRB had a different understanding of federal regulations and research ethics than did the OHRP—a result most likely explained by vague and unspecific regulatory requirements. One example is the lack of clarity surrounding the concept of "research involving human subjects." If an IRB has one interpretation and a regulatory office another, both research and advances in practice can be held up and can even lead to disastrous federal penalties if the wrong judgment is made.

In the Pronovost case, the activities involved no new interventions and posed no risk for patients. Research was fully integrated with practice, and physicians were following the safest practices known to exist—without introducing new research activities. OHRP officials made the judgment that because infection rates were being studied in *patients*, the study called for full committee review and for the informed consent of *subjects*. But this

research was by its design an attempt to improve medical care. The invocation of regulations intended to protect research subjects led to a delay in the use of effective preventive measures in hospitals that may have caused multiple patient deaths and could have eventuated in unjustified penalties to the medical research institutions and hospitals involved.

Eventually the OHRP issued a statement that in effect admitted that it had been wrong. It acknowledged that the work was "being used … solely for clinical purposes, not medical research or experimentation." The OHRP further acknowledged that the activity, from the start, "would likely have been eligible for both expedited IRB review and a waiver of the informed consent requirement."[112] While laudable, this acknowledgment of error is puzzling. Pronovost's work was an empirical study and therefore research. Perhaps the OHRP means that the study is research, though not "research involving human subjects." This estimate is probably the correct judgment, but it also indicates that the notion of research involving human subjects is systematically unclear, which can lead to overprotection, as in this case, thus causing harm.

Government regulations usually need some form of interpretation, but we should not tolerate a system in which lives might be lost because of an obsolete conception of human-subjects research that obstructs riskless studies aimed at improving medical practice. When research investigations are unduly restricted through requirements of regulation and review, the requirements should be adjusted. In the case of Pronovost's research, the initial IRB review was correct when it concluded that the study did not need full IRB review and patients' informed consent, but later the system of oversight worked more to present risks to current and future patients than to protect them.

Problems of Group Harm in Research

In Chapter 4 (pp. 119–23), we presented a theory of valid informed consent. In addition to the paradigmatic case of specific, explicit informed consent, we also examined the place of other varieties of consent, including general, implicit, tacit, and presumed consent. We now turn to a version of "general consent," often called "broad consent," "global consent," or "blanket consent," in the context of research using biological samples. Under this form of consent, harms may occur for individuals and groups as a result of inadequate information and understanding. The problems can be acute when biological samples are banked and subsequently used in unanticipated ways that may harm individuals or groups. Valid informed consent is one protective measure, but it is insufficient by itself. Improved forms of governance of banks of biological specimens are also needed.[113]

Research on stored biological specimens. Advances in science have introduced confusion about how we can efficiently promote research while protecting the rights of donors of samples. Samples collected for future research may not be adequately described in a protocol or consent form when the collection occurs. The wording in the form may be dictated by shadowy anticipated future uses of samples, with little explanation of possible harmful outcomes. The challenge is not to cause harm to personal and group interests and not to violate privacy and confidentiality. The moral problem is whether it is possible to meet this challenge and, if so, how.[114]

Samples and data frequently descend from sources external to a research setting, including industry, government, and university sources, and it may be difficult to determine both whether adequately informed consent was obtained for use of the samples and data and whose interests might be at risk. Using samples or data to achieve goals other than those initially disclosed to subjects negates even an originally valid consent process and threatens the trust between subjects and investigators. Even anonymized samples can harm some personal and group interests and may violate the investigator-subject relationship. Furthermore, secure anonymization is notoriously difficult to achieve, as various breaches of privacy have shown.

We will not try to resolve all of these complicated issues. We will instead present a paradigm case that exemplifies the pitfalls and risks of harm in research that permits broad consents.

Diabetes research on Havasupai Indians. This case involves research conducted at Arizona State University using as research subjects the Havasupai Indians of the Grand Canyon. Investigators used a broad consent, which was not as carefully scrutinized by university ethics committee review as it should have been. The story

starts in 1990 when members of the fast-disappearing Havasupai tribe gave DNA samples to university researchers with the goal of providing genetic information about the tribe's distressing, indeed alarming, rate of diabetes. Beginning in the 1960s, the Havasupai had experienced a high incidence of type 2 diabetes that led to amputations and forced many tribal members to leave their village in the Grand Canyon to live closer to dialysis centers.

From 1990 to 1994, approximately one hundred members of the tribe signed an Arizona State University broad consent that stated the research was to "study the causes of behavioral/medical disorders." The consent form was intentionally confined to clear, simply written, basic information, because English is a second language for many Havasupai, and few of the tribe's remaining 650 members had graduated from high school. From the researchers' perspective, tribe members had consented to collection of blood and to its use in genetic research well beyond the research on their particular disease. The Havasupai subjects, by contrast, denied that they gave permission for any nondiabetes research and insisted that they received inadequate information about and had an inadequate understanding of the risks of the research before they agreed to participate.

In the course of the research, diabetes was investigated, but the roughly two hundred blood samples were also put to several additional uses in genetics research having nothing to do with diabetes. One use was to study mental illness, especially schizophrenia, and another was to examine inbreeding in the tribe. Approximately two dozen scholarly articles were published on the basis of research on the samples. To the Havasupai, some of this research was offensive, insulting, stigmatizing, and harmful, and also a provocative examination of taboo subjects. They filed a lawsuit charging research investigators with a failure to obtain informed consent, unapproved use of data, infliction of emotional distress, and violation of medical confidentiality. Charges included fraud, breach of fiduciary duty, negligence, violation of civil rights, and trespass.[115]

Both the researchers and the review committee at the university apparently did not notice the serious risks of harm, disrespect, and abuse inherent in the research they conducted subsequent to the broad consent. One article eventually published by investigators theorized that the tribe's ancestors had crossed the frozen Bering Sea to arrive in North America. This thesis directly contradicted the tribe's traditional stories and cosmology, which have quasi-religious significance for the tribe. According to its tradition, the tribe originated in the Grand Canyon and was assigned to be the canyon's guardian. It was to them disorienting and abhorrent to be told that the tribe was instead probably of Asian origin and that this hypothesis was developed from studies of their blood, which also has a special significance to the Havasupai. The thesis also set off legal alarms in the community, because the Havasupai had previously argued that their origin in the Grand Canyon was the legal basis of their entitlement to the land. The National Congress of American Indians has pointed out that many native American tribes are in conditions of vulnerability similar to those of the Havasupai.[116]

This case presents paradigmatic problems of risk of harm, inadequate consent, and violations of human rights. In particular, it underlines the need to attend to group, as well as individual, harms, and to a richer conception of harms in research than often occurs. Research on samples, especially genetics research, can create psychosocial risks in the absence of physical risks to individual sources of the samples. In this case the tribe was harmed by the damage to its traditional self-understanding. This case also raises questions about whether scientists took advantage of a vulnerable population by exploiting its members' lack of understanding.

In the end, the university made a compensatory payment of $700,000 to the affected tribal members, provided funds for a school and clinic, and returned the DNA samples. The university acknowledged that the total compensation package was to "remedy the wrong that was done."[117] The university had worked for years to establish good relationships with Native American tribes in Arizona, but this reservoir of trust was profoundly damaged by these events.

CONCLUSION

We have concentrated in this chapter on the principle of nonmaleficence and its implications for refusals of treatment and requests for assistance in dying when the patient's death is highly probable or certain or when the

patient's quality of life is very poor, and on its implications for the protection of individuals and groups from harm in the clinic and in research. From the principle that we should avoid causing harm to persons, there is no direct step to the conclusion that a positive obligation exists to provide benefits such as health care and various forms of assistance. We have not entered this territory in this chapter on nonmaleficence because obligations to provide positive benefits are the territory of beneficence and justice. We treat these principles in Chapters 6 and 7.

NOTES

1. 1. W. H. S. Jones, *Hippocrates*, vol. I (Cambridge, MA: Harvard University Press, 1923), p. 165. See also Albert R. Jonsen, "Do No Harm: Axiom of Medical Ethics," in *Philosophical and Medical Ethics: Its Nature and Significance*, ed. Stuart F. Spicker and H. Tristram Engelhardt, Jr. (Dordrecht, Netherlands: D. Reidel, 1977), pp. 27–41; and Steven H. Miles, *The Hippocratic Oath and the Ethics of Medicine* (New York: Oxford University Press, 2004).
2. 2. W. D. Ross, *The Right and the Good* (Oxford: Clarendon, 1930), pp. 21–26; John Rawls, *A Theory of Justice* (Cambridge, MA: Harvard University Press, 1971; rev. ed., 1999), p. 114 (1999: p. 98).
3. 3. William Frankena, *Ethics*, 2nd ed. (Englewood Cliffs, NJ: Prentice Hall, 1973), p. 47.
4. 4. On the idea that there is a priority of avoiding harm, see criticisms by N. Ann Davis, "The Priority of Avoiding Harm," in *Killing and Letting Die*, 2nd ed., ed. Bonnie Steinbock and Alastair Norcross (New York: Fordham University Press, 1999), pp. 298–354.
5. 5. Bernard Gert presents a theory of this sort. He accepts numerous obligations of nonmaleficence while holding that beneficence is entirely in the realm of moral ideals, not the realm of obligations. See our interpretation and critique of his theory in Chapter 10, pp. 428–32.
6. 6. *McFall v. Shimp*, no. 78-1771 in Equity (C. P. Allegheny County, PA, July 26, 1978); Barbara J. Culliton, "Court Upholds Refusal to Be Medical Good Samaritan," *Science* 201 (August 18, 1978): 596–97; Mark F. Anderson, "Encouraging Bone Marrow Transplants from Unrelated Donors," *University of Pittsburgh Law Review* 54 (1993): 477ff.
7. 7. Alan Meisel and Loren H. Roth, "Must a Man Be His Cousin's Keeper?" *Hastings Center Report* 8 (October 1978): 5–6. For further analysis of this case, see Guido Calabresi, "Do We Own Our Bodies?" *Health Matrix* 1 (1991): 5-18, available at Faculty Scholarship Series. Paper 2011, Yale Law School Legal Scholarship Repository, available at http://digitalcommons.law.yale.edu/fss_papers/2011 (accessed September 4, 2018).
8. 8. Joel Feinberg, *Harm to Others*, vol. I of *The Moral Limits of the Criminal Law* (New York: Oxford University Press, 1984), pp. 32–36, and also 51–55, 77–78.
9. 9. The best definition of *harm* is philosophically controversial. For different accounts that would modify our definition (which is indebted to Feinberg), see Elizabeth Harman, "Harming as Causing Harm," in *Harming Future Persons*, ed. Melinda Roberts and David Wasserman (New York: Springer, 2009), pp. 137–54; Seana Shiffrin, "Wrongful Life, Procreative Responsibility, and the Significance of Harm," *Legal Theory* 5 (1999): 117–48; and Alastair Norcross, "Harming in Context," *Philosophical Studies* 123 (2005): 149–73.
10. 10. On some of the many roles of harm and nonmaleficence in bioethics, see Bettina Schöne-Seifert, "Harm," in *Bioethics* (formerly *Encyclopedia of Bioethics*), 4th ed., ed. Bruce Jennings (Farmington Hills, MI: Gale, Cengage Learning, 2014), vol. 3, pp. 1381–86.
11. 11. For an interesting account of the central rules of nonmaleficence and their role in bioethics, see Bernard Gert, *Morality: Its Nature and Justification* (New York: Oxford University Press, 2005); and Gert, Charles M. Culver, and K. Danner Clouser, *Bioethics: A Systematic Approach* (New York: Oxford University Press, 2006).
12. 12. H. L. A. Hart, *Punishment and Responsibility* (Oxford: Clarendon, 1968), esp. pp. 136–57; Joel Feinberg, *Doing and Deserving* (Princeton, NJ: Princeton University Press, 1970), esp. pp. 187–221; Eric D'Arcy, *Human Acts: An Essay in Their Moral Evaluation* (Oxford: Clarendon, 1963), esp. p. 121. For a revealing empirical study useful for biomedical ethics, see A. Russell Localio, Ann G. Lawthers, Troyen A. Brennan, et al., "Relation between Malpractice Claims and Adverse Events Due to Negligence—

Results of the Harvard Medical Practice Study III," *New England Journal of Medicine* 325 (1991): 245–51.

13. 13. On medical negligence, medical error, physician-caused harm, and their connection to medical ethics, see Virginia A. Sharpe and Alan I. Faden, *Medical Harm: Historical, Conceptual, and Ethical Dimensions of Iatrogenic Illness* (New York: Cambridge University Press, 1998); and Milos Jenicek, *Medical Error and Harm: Understanding, Prevention, and Control* (New York: CRC Press/Productivity Press of Taylor & Francis, 2011). See also R. C. Solomon, "Ethical Issues in Medical Malpractice," *Emergency Medicine Clinics of North America* 24, no. 3 (2006): 733–47.
14. 14. As quoted in Angela Roddy Holder, *Medical Malpractice Law* (New York: Wiley, 1975), p. 42.
15. 15. Cf. the conclusions about physicians' reservations in Arthur R. Derse, "Limitation of Treatment at the End-of-Life: Withholding and Withdrawal," *Clinics in Geriatric Medicine* 21 (2005): 223–38; Neil J. Farber et al., "Physicians' Decisions to Withhold and Withdraw Life-Sustaining Treatments," *Archives of Internal Medicine* 166 (2006): 560–65; and Sharon Reynolds, Andrew B. Cooper, and Martin McKneally, "Withdrawing Life-Sustaining Treatment: Ethical Considerations," *Surgical Clinics of North America* 87 (2007): 919–36, esp. 920–23. For a comprehensive examination of medical ethics issues that have arisen about this distinction in the British context, see Medical Ethics Department, British Medical Association, *Withholding and Withdrawing Life-prolonging Medical Treatment: Guidance for Decision Making*, 3rd ed. (Oxford: BMJ Books, Blackwell, John Wiley, 2007).
16. 16. The long-standing distinction between "extraordinary" or "heroic" and "ordinary" means of treatment still sometimes appears in popular discourse, as in this case. It has had a long history, particularly in Roman Catholic moral theology and philosophy where refusing "ordinary" treatment constituted a suicide and withholding or withdrawing "ordinary" treatment constituted a homicide. By contrast, refusing or withholding/withdrawing "extraordinary" treatment could be morally justified in various circumstances. This distinction has now been largely abandoned because the terms became attached to usual and unusual or customary and uncustomary treatments, without regard to the balance of benefits and burdens for the patients receiving those treatments, and proponents of the distinction developed a variety of other morally irrelevant criteria, such as simple and complex, to explicate these notions. In Roman Catholic thought, the common replacement terms are "proportionate" and "disproportionate." See, for example, the United States Conference of Catholic Bishops (USCB), *Ethical and Religious Directives for Catholic Health Services*, 6th ed. (Washington, DC: USCB, issued June 2018), Part 5, available at http://www.usccb.org/about/doctrine/ethical-and-religious-directives/upload/ethical-religious-directives-catholic-health-service-sixth-edition-2016-06.pdf (accessed September 11, 2018). On the nature and evolution of the doctrine in Roman Catholic thought, see Scott M. Sullivan, "The Development and Nature of the Ordinary/Extraordinary Means Distinction in the Roman Catholic Tradition," *Bioethics* 21 (2007): 386–97; Donald E. Henke, "A History of Ordinary and Extraordinary Means," *National Catholic Bioethics Quarterly* 5 (2005): 555–75; and Kevin W. Wildes, "Ordinary and Extraordinary Means and the Quality of Life," *Theological Studies* 57 (1996): 500–512. See also Jos V. M. Welie, "When Medical Treatment Is No Longer in Order: Toward a New Interpretation of the Ordinary-Extraordinary Distinction," *National Catholic Bioethics Quarterly* 5 (2005): 517–36.
17. 17. This case was presented to one of the authors during a consultation.
18. 18. For defenses of the distinction along these or similar lines, see Daniel P. Sulmasy and Jeremy Sugarman, "Are Withholding and Withdrawing Therapy Always Morally Equivalent?" *Journal of Medical Ethics* 20 (1994): 218–22 (commented on by John Harris, pp. 223–24); and Kenneth V. Iserson, "Withholding and Withdrawing Medical Treatment: An Emergency Medicine Perspective," *Annals of Emergency Medicine* 28 (1996): 51–54. For opposing positions on the moral equivalence of withholding and withdrawing, see Lars Øystein Ursin, "Withholding and Withdrawing Life-Sustaining Treatment: Ethically Equivalent?" *American Journal of Bioethics* 19 (2019): 10–20; and Dominic Wilkinson, Ella Butcherine, and Julian Savulescu, "Withdrawal Aversion and the Equivalence Test," *American Journal of Bioethics* 19 (2019): 21–28, followed by several commentaries.
19. 19. *In the matter of Spring*, Mass. 405 N.E. 2d 115 (1980), at 488–89.
20. 20. Lewis Cohen, Michael Germain, and David Poppel, "Practical Considerations in Dialysis Withdrawal," *JAMA: Journal of the American Medical Association* 289 (2003): 2113–19. A study of a French population receiving dialysis found that 20.4% of patients "died following withdrawal from dialysis": Béatrice Birmelé, Maud François, Josette Pengloan, et al., "Death after Withdrawal from

Dialysis: The Most Common Cause of Death in a French Dialysis Population," *Nephrology Dialysis Transplantation* 19 (2004): 686–91. The authors hold that discontinuation of dialysis is a more common cause of death in patients in North America and the United Kingdom than in "the rest of Europe." A retrospective study in Australia and New Zealand found that dialysis withdrawal accounted for more than one in four deaths among patients with end-stage renal disease in the period 1999–2008. See Hoi Wong Chan et al., "Risk Factors for Dialysis Withdrawal: An Analysis of the Australia and New Zealand Transplant (ANZDATA) Registry, 1999–2008)," *Clinical Journal of the American Society of Nephrology* 7, no. 5 (May 7, 2012): 775–81. Some studies, but not all, distinguish death caused by dialysis withdrawal and death caused by the illness that led to the dialysis withdrawal. See Milagros Ortiz et al., "Dialysis Withdrawal: Cause of Mortality along a Decade (2004–2014)," *Nephrology, Dialysis, Transplantation* 32, issue supplement 3 (May 26, 2017): iii358–iii359.

21. 21. See Rebecca J. Schmidt and Alvin H. Moss, "Dying on Dialysis: The Case for a Dignified Withdrawal," *Clinical Journal of the American Society of Nephrology* 9, no. 1 (2014): 174–80.
22. 22. Robert Stinson and Peggy Stinson, *The Long Dying of Baby Andrew* (Boston: Little, Brown, 1983), p. 355.
23. 23. Katy Butler, "What Broke My Father's Heart," *New York Times Magazine*, June 18, 2010, available at http://www.nytimes.com/2010/06/20/magazine/20pacemaker-t.html?pagewanted=all (accessed July 4, 2018). A fuller version of the story appears in Butler, *Knocking on Heaven's Door: The Path to a Better Way of Death* (New York: Scribner, 2013). For clinicians' views and ethical analyses, see Michael B. Bevins, "The Ethics of Pacemaker Deactivation in Terminally Ill Patients," *Journal of Pain and Symptom Management* 41 (June 2011): 1106–10; T. C. Braun et al., "Cardiac Pacemakers and Implantable Defibrillators in Terminal Care," *Journal of Pain and Symptom Management* 18 (1999): 126–31; Daniel B. Kramer, Susan L. Mitchell, and Dan W. Brock, "Deactivation of Pacemakers and Implantable Cardioverter-Defibrillators," *Progress in Cardiovascular Diseases* 55, no. 3 (November–December 2012): 290–99; and K. E. Karches and D. P. Sulmasy, "Ethical Considerations for Turning Off Pacemakers and Defibrillators," *Cardiac Electrophysiology Clinics* 7, no. 3 (September 2015): 547–55.
24. 24. Paul Mueller et al., "Deactivating Implanted Cardiac Devices in Terminally Ill Patients: Practices and Attitudes," *Pacing and Clinical Electrophysiology* 31, no. 5 (2008): 560–68. See also the study reported by Daniel B. Kramer, Aaron S. Kesselheim, Dan W. Brock, and William H. Maisel, "Ethical and Legal Views of Physicians Regarding Deactivation of Cardiac Implantable Electric Devices: A Quantitative Assessment," *Heart Rhythm* 7, no. 11 (November 2010): 1537–42; and A. S. Kelley et al., "Implantable Cardioverter-Defibrillator Deactivation at End-of-Life: A Physician Survey," *American Heart Journal* 157 (2009): 702–8. For nurses' concerns and general support for deactivation of cardiovascular implantable electronic devices, see D. B. Kramer et al., "'Just Because We Can Doesn't Mean We Should': Views of Nurses on Deactivation of Pacemakers and Implantable Cardioverter-Defibrillators," *Journal of Interventional Cardiac Electrophysiology* 32, no. 3 (December 2011): 243–52.
25. 25. Rachel Lampert et al., "HRS Expert Consensus Statement on the Management of Cardiovascular Implantable Electronic Devices (CIEDs) in Patients Nearing End of Life or Requesting Withdrawal of Therapy," *Heart Rhythm* 7, no. 7 (July 2010): 1008–25, available at https://www.heartrhythmjournal.com/article/S1547-5271(10)00408-X/abstract (accessed July 4, 2018).
26. 26. Lampert et al., "HRS Expert Consensus Statement on the Management of Cardiovascular Implantable Electronic Devices (CIEDs)."
27. 27. Mueller et al., "Deactivating Implanted Cardiac Devices in Terminally Ill Patients: Practices and Attitudes," p. 560. More attention needs to be paid to the role and responsibility of the industry representative in deactivating these devices.
28. 28. See Jeffrey P. Burns and Robert D. Truog, "The DNR Order after 40 Years," *New England Journal of Medicine* 375 (August 11, 2016): 504–6; Susanna E. Bedell and Thomas L. Delbanco, "Choices about Cardiopulmonary Resuscitation in the Hospital: When Do Physicians Talk with Patients?" *New England Journal of Medicine* 310 (April 26, 1984): 1089–93; and Marcia Angell, "Respecting the Autonomy of Competent Patients," *New England Journal of Medicine* 310 (April 26, 1984): 1115–16. In one survey, 50% of the physicians responding opposed unilateral DNR orders; physicians supporting such orders were more likely to be in pulmonary/critical care medicine. See Michael S. Putnam et al., "Unilateral Do Not Resuscitate Orders: Physician Attitudes and Practices," *Chest* 152, no. 1 (July 2017): 224–25.

29. 29. See Evie G. Marcolini, Andrew T. Putnam, and Ani Aydin, "History and Perspectives on Nutrition and Hydration at the End of Life," *Yale Journal of Biology and Medicine* 91, no. 2 (June 2018): 173–76. They write: "ANH are defined as a group of medical treatments provided to patients who cannot meet their daily requirements orally, with resultant malnutrition, electrolyte abnormalities, and/or metabolic derangements. The various modalities to deliver ANH include intravenous hydration and intravenous parenteral nutrition, nasogastric feeding, and placement of surgical feeding devices to deliver the required hydration and nourishment."
30. 30. See Joanne Lynn and James F. Childress, "Must Patients Always Be Given Food and Water?" *Hastings Center Report* 13 (October 1983): 17–21; reprinted in *By No Extraordinary Means: The Choice to Forgo Life-Sustaining Food and Water*, ed. Joanne Lynn (Bloomington: Indiana University Press, 1986, expanded edition, 1989), pp. 47–60; and Childress, "When Is It Morally Justifiable to Discontinue Medical Nutrition and Hydration?" in *By No Extraordinary Means*, ed. Lynn, pp. 67–83.
31. 31. This case has been adapted with permission from a case presented by Dr. Martin P. Albert of Charlottesville, Virginia. On problems in nursing homes, see Alan Meisel, "Barriers to Forgoing Nutrition and Hydration in Nursing Homes," *American Journal of Law and Medicine* 21 (1995): 335–82; and Sylvia Kuo et al., "Natural History of Feeding-Tube Use in Nursing Home Residents with Advanced Dementia," *Journal of the American Medical Directors Association* 10 (2009): 264–70, which concludes that most feeding tubes are inserted during an acute care hospitalization and are associated with poor survival and subsequent heavy use of health care. O'Brien and colleagues determined that close to 70% of nursing home residents prefer not to have a feeding tube placed in cases of permanent brain damage, and many others shared that preference when they learned that physical restraints might be required: Linda A. O'Brien et al., "Tube Feeding Preferences among Nursing Home Residents," *Journal of General Internal Medicine* 12 (1997): 364–71. In line with research that has indicated little benefit coupled with unnecessary suffering, the insertion of feeding tubes in US nursing home residents with advanced dementia declined substantially from 2000 to 2014: from 12% to 6%, with higher rates of use by black than white residents. See Susan L. Mitchell et al., "Tube Feeding in US Nursing Home Residents with Advanced Dementia, 2000–2014," *JAMA: Journal of the American Medical Association* 316, no. 7 (2016): 769–70.
32. 32. *In the matter of Quinlan*, 70 N.J. 10, 355 A.2d 647, cert. denied, 429 U.S. 922 (1976). The New Jersey Supreme Court ruled that the Quinlans could disconnect the mechanical ventilator so that the patient could "die with dignity."
33. 33. See Joseph Quinlan, Julia Quinlan, and Phyllis Battell, *Karen Ann: The Quinlans Tell Their Story* (Garden City, NY: Doubleday, 1977).
34. 34. In *Cruzan v. Director, Missouri Dep't of Health*, 497 U.S. 261 (1990), the US Supreme Court concluded that a competent person has a constitutionally protected right to refuse lifesaving hydration and nutrition. Its dicta reflected no distinction between medical and sustenance treatments.
35. 35. See Lois Shepherd, *If That Ever Happens to Me: Making Life and Death Decisions after Terri Schiavo* (Chapel Hill: University of North Carolina Press, 2009); Timothy E. Quill, "Terri Schiavo—A Tragedy Compounded," *New England Journal of Medicine* 352, no. 16 (2005): 1630–33; George J. Annas, "'Culture of Life' Politics at the Bedside—The Case of Terri Schiavo," *New England Journal of Medicine* 352, no. 16 (2005): 1710–15; and Tom Koch, "The Challenge of Terri Schiavo: Lessons for Bioethics," *Journal of Medical Ethics* 31 (2005): 376–78). See further Thomas S. Shannon, "Nutrition and Hydration: An Analysis of the Recent Papal Statement in the Light of the Roman Catholic Bioethical Tradition," *Christian Bioethics* 12 (2006): 29–41.
36. 36. M. I. Del Rio et al., "Hydration and Nutrition at the End of Life: A Systematic Review of Emotional Impact, Perceptions, and Decision-Making among Patients, Family, and Health Care Staff," *Psycho-oncology* 21, no. 9 (September 2012): 913–21.
37. 37. See C. M. Callahan et al., "Decision-making for Percutaneous Endoscopic Gastrotomy among Older Adults in a Community Setting," *Journal of the American Geriatrics Society* 47 (1999): 1105–9.
38. 38. For a summary of the available evidence, see Howard Brody et al., "Artificial Nutrition and Hydration: The Evolution of Ethics, Evidence, and Policy," *Journal of General Internal Medicine* 26, no. 9 (2011): 1053–58.
39. 39. The RDE has rough precedents that predate the writings of St. Thomas Aquinas (e.g., in St. Augustine and Abelard). However, the history primarily flows from St. Thomas. See Anthony Kenny, "The History

of Intention in Ethics," in *Anatomy of the Soul* (Oxford: Basil Blackwell, 1973), Appendix; Joseph T. Mangan, "An Historical Analysis of the Principle of Double Effect," *Theological Studies* 10 (1949): 41–61; and T. A. Cavanaugh, *Double-Effect Reasoning: Doing Good and Avoiding Evil* (New York: Oxford University Press, 2006), chap. 1.

40. 40. For an overview of the doctrine of double effect, see Alison McIntyre, "Doctrine of Double Effect," *The Stanford Encyclopedia of Philosophy* (Winter 2014 Edition), ed. Edward N. Zalta, available at https://plato.stanford.edu/archives/win2014/entries/double-effect/ (accessed June 28, 2018); Suzanne Uniacke, "The Doctrine of Double Effect," in *Principles of Health Care Ethics*, 2nd ed., ed. Richard E. Ashcroft et al. (Chichester, England: John Wiley, 2007), pp. 263–68. For several representative philosophical positions, see P. A. Woodward, ed., *The Doctrine of Double Effect: Philosophers Debate a Controversial Moral Principle* (Notre Dame, IN: Notre Dame University Press, 2001). In an influential interpretation, Joseph Boyle reduces the RDE to two conditions: intention and proportionality. "Who Is Entitled to Double Effect?" *Journal of Medicine and Philosophy* 16 (1991): 475–94; and "Toward Understanding the Principle of Double Effect," *Ethics* 90 (1980): 527–38.

    For criticisms of intention-weighted views, see Timothy E. Quill, Rebecca Dresser, and Dan Brock, "The Rule of Double Effect—A Critique of Its Role in End-of-Life Decision Making," *New England Journal of Medicine* 337 (1997): 1768–71; Alison MacIntyre, "Doing Away with Double Effect," *Ethics* 111, no. 2 (2001): 219–55; and Sophie Botros, "An Error about the Doctrine of Double Effect," *Philosophy* 74 (1999): 71–83. T. M. Scanlon rejects the RDE on the grounds that it is not clear how an agent's intentions determine the permissibility of an agent's actions, as the doctrine claims; however, it may still be appropriate in assessing the reasons an agent saw as bearing on his actions. Scanlon, *Moral Dimensions: Permissibility, Meaning, Blame* (Cambridge, MA: Harvard University Press, 2008), esp. Introduction and chaps. 1–2.

41. 41. For assessments, see Daniel Sulmasy, "Reinventing the Rule of Double Effect," in *The Oxford Handbook of Bioethics*, ed. Bonnie Steinbock (New York: Oxford University Press, 2010), pp. 114–49; David Granfield, *The Abortion Decision* (Garden City, NY: Image Books, 1971); and Susan Nicholson, *Abortion and the Roman Catholic Church* (Knoxville, TN: Religious Ethics, 1978). See also the criticisms of the RDE in Donald Marquis, "Four Versions of Double Effect," *Journal of Medicine and Philosophy* 16 (1991): 515–44, reprinted in *The Doctrine of Double Effect*, ed. Woodward, pp. 156–85.
42. 42. See Michael Bratman, *Intention, Plans, and Practical Reason* (Cambridge, MA: Harvard University Press, 1987).
43. 43. Alvin I. Goldman, *A Theory of Human Action* (Englewood Cliffs, NJ: Prentice Hall, 1970), pp. 49–85.
44. 44. See the analysis in Hector-Neri Castañeda, "Intensionality and Identity in Human Action and Philosophical Method," *Nous* 13 (1979): 235–60, esp. 255.
45. 45. Our analysis here draws from Ruth R. Faden and Tom L. Beauchamp, *A History and Theory of Informed Consent* (New York: Oxford University Press, 1986), chap. 7.
46. 46. We also follow John Searle in thinking that we cannot reliably distinguish, in many situations, between acts, effects, consequences, and events. Searle, "The Intentionality of Intention and Action," *Cognitive Science* 4 (1980): 65.
47. 47. This interpretation of double effect is defended by Boyle, "Who Is Entitled to Double Effect?"
48. 48. See the arguments in Joseph Boyle, "Medical Ethics and Double Effect: The Case of Terminal Sedation," *Theoretical Medicine* 25 (2004): 51–60; Boyle, "The Relevance of Double Effect to Decisions about Sedation at the End of Life," in *Sedation at the End-of-Life: An Interdisciplinary Approach*, ed. Paulina Taboada (Dordrecht: Springer Science+Business Media, 2015), pp. 55–72; Alejandro Miranda, "The Field of Application of the Principle of the Double Effect and the Problem of Palliative Sedation," in *Sedation at the End-of-Life*, ed. Taboada, pp. 73–90; Kasper Raus, Sigrid Sterckx, and Freddy Mortier, "Can the Doctrine of Double Effect Justify Continuous Deep Sedation at the End of Life?" in *Continuous Sedation at the End of Life: Ethical, Clinical and Legal Perspectives*, ed. Sigrid Sterckx and Kasper Raus (Cambridge: Cambridge University Press, 2017), pp. 177–201; Alison McIntyre, "The Double Life of Double Effect," *Theoretical Medicine and Bioethics* 25 (2004): 61–74; Daniel P. Sulmasy and Edmund D. Pellegrino, "The Rule of Double Effect: Clearing Up the Double Talk," *Archives of Internal Medicine* 159 (1999): 545–50; Lynn A. Jansen and Daniel Sulmasy, "Sedation, Alimentation, Hydration, and

Equivocation: Careful Conversation about Care at the End of Life," *Annals of Internal Medicine* 136 (June 4, 2002): 845–49; and Johannes J. M. van Delden, "Terminal Sedation: Source of a Restless Ethical Debate," *Journal of Medical Ethics* 33 (2007): 187–88.

49. 49. See Quill, Dresser, and Brock, "The Rule of Double Effect"; and McIntyre, "The Double Life of Double Effect."
50. 50. Lawrence Masek, "Intention, Motives, and the Doctrine of Double Effect," *Philosophical Quarterly* 60, no. 240 (July 2010): 567–85, which argues that "the moral permissibility of an action depends at least partly on how it forms an agent's character." See also Masek, *Intention, Character, and Double Effect* (Notre Dame, IN: University of Notre Dame Press, 2018).
51. 51. Debates about the proper analysis of the concept of medical futility have been vigorous over the last few decades. See Dominic James Wilkinson and Julian Savulescu, "Knowing When to Stop: Futility in the Intensive Care Unit," *Current Opinion in Anesthesiology* 24 (April 2011): 160–65; Ben White, Lindy Willmott, Eliana Close, et al., "What Does 'Futility' Mean? An Empirical Study of Doctors' Perceptions," *Medical Journal of Australia* 204 (2016), available online at https://www.mja.com.au/journal/2016/204/8/what-does-futility-mean-empirical-study-doctors-perceptions (accessed June 29, 2018); James L. Bernat, "Medical Futility: Definition, Determination, and Disputes in Critical Care," *Neurocritical Care* 2 (2005): 198–205; D. K. Sokol, "The Slipperiness of Futility," *BMJ: British Medical Journal* 338 (June 5, 2009); E. Chwang, "Futility Clarified," *Journal of Law, Medicine, & Ethics* 37 (2009): 487–95; Baruch A. Brody and Amir Halevy, "Is Futility a Futile Concept?" *Journal of Medicine and Philosophy* 20 (1995): 123–44; R. Lofmark and T. Nilstun, "Conditions and Consequences of Medical Futility," *Journal of Medical Ethics* 28 (2002): 115–19; and Loretta M. Kopelman, "Conceptual and Moral Disputes about Futile and Useful Treatments," *Journal of Medicine and Philosophy* 20 (1995): 109–21. Important books in the debate include Susan B. Rubin, *When Doctors Say No: The Battleground of Medical Futility* (Bloomington: Indiana University Press, 1998); and Lawrence J. Schneiderman and Nancy S. Jecker, *Wrong Medicine: Doctors, Patients, and Futile Treatment*, 2nd ed. (Baltimore: Johns Hopkins University Press, 2011). A cross-national view of values, policies, and practices appears in Alireza Bagheri, ed., *Medical Futility: A Cross-National Study* (London: Imperial College Press, 2013)
52. 52. See Wilkinson and Savulescu, "Knowing When to Stop," which proposes the language of "medically inappropriate" to highlight that medical professionals are making value judgments and that an intervention is appropriate or inappropriate for realizing some goal of treatment. For a discussion of the limits of providing requested "nonbeneficial interventions," see Allan S. Brett and Laurence B. McCullough, "Addressing Requests by Patients for Nonbeneficial Interventions," *JAMA: Journal of the American Medical Association* 307 (January 11, 2012): 149–50.
53. 53. G. T. Bosslet et al., "An Official ATS/AACN/ACCP/ESICM/SCCM Policy Statement: Responding to Requests for Potentially Inappropriate Treatments in Intensive Care Units," *American Journal of Respiratory Critical Care Medicine* 191, no. 11 (2015): 1318–30; J. L. Nates et al., "ICU Admission, Discharge, and Triage Guidelines: A Framework to Enhance Clinical Operations, Development of Institutional Policies, and Further Research,"*Critical Care Medicine* 44, no. 8 (2016): 1553–1602.
54. 54. Bosslett et al., "An Official ATS/AACN/ACCP/ESICM/SCCM Policy Statement: Responding to Requests for Potentially Inappropriate Treatments in Intensive Care Units," p. 1318.
55. 55. In a special issue of *Perspectives in Biology and Medicine* 60, no. 3 (Summer 2017) devoted to futility, Lawrence J. Schneiderman, Nancy S. Jecker, and Albert R. Jonsen's "The Abuse of Futility," responds to critiques of medical futility and to efforts to develop conceptions of "inappropriate" treatment. In response to this lead article, twenty-one additional articles address these issues.
56. 56. For a defense of an occasional compassionate futile intervention, see Robert D. Truog, "Is It Always Wrong to Perform Futile CPR?" *New England Journal of Medicine* 362 (2010): 477–79. A counterargument, based on the individual's right to die with dignity, appears in J. J. Paris, P. Angelos, and M. D. Schreiber, "Does Compassion for a Family Justify Providing Futile CPR?" *Journal of Perinatology* 30 (December 2010): 770–72.
57. 57. See further John Luce, "A History of Resolving Conflicts over End-of-Life Care in Intensive Care Units in the United States," *Critical Care Medicine* 38 (August 2010): 1623–29. For constructive proposals that take account of legitimate disagreement, see Amir Halevy and Baruch A. Brody, "A Multi-Institution Collaborative Policy on Medical Futility," *JAMA: Journal of the American Medical Association*

276 (1996): 571–75; and Carolyn Standley and Bryan A. Liang, "Addressing Inappropriate Care Provision at the End-of-Life: A Policy Proposal for Hospitals," *Michigan State University Journal of Medicine and Law* 15 (Winter 2011): 137–76. Since 1999, the Texas Advance Directives Act, sometimes erroneously referred to as the "Texas Futile Care Law," has allowed physicians under certain conditions to unilaterally discontinue life-sustaining treatments deemed futile, after giving notice and waiting ten days. See the following discussions: Robert L. Fine, "Point: The Texas Advance Directives Act Effectively and Ethically Resolves Disputes about Medical Futility," *Chest* 136 (2009): 963–67; Robert D. Truog, "Counterpoint: The Texas Advance Directives Act Is Ethically Flawed: Medical Futility Disputes Must Be Resolved by a Fair Process," *Chest* 136 (2009): 968–71, followed by discussion 971–73; Wilkinson and Savulescu, "Knowing When to Stop"; and Robert M. Veatch, "So-Called Futile Care: The Experience of the United States," in *Medical Futility: A Cross-National Study*, ed. Bagheri, pp. 24–28. For a proposal to retain the option in "futility" cases of appeal to the courts because of its benefits at the societal level, see Douglas B. White and Thaddeus M. Pope, "The Courts, Futility, and the Ends of Medicine," *JAMA: Journal of the American Medical Association* 307 (2012): 151–52.

58. 58. *Superintendent of Belchertown State School v. Saikewicz*, Mass., 370 N.E. 2d 417 (1977), at 428.
59. 59. Paul Ramsey, *Ethics at the Edges of Life: Medical and Legal Intersections* (New Haven, CT: Yale University Press, 1978), p. 155.
60. 60. See President's Commission for the Study of Ethical Problems in Medicine and Behavioral Research, *Deciding to Forego Life-Sustaining Treatment: Ethical, Medical, and Legal Issues in Treatment Decisions* (Washington, DC: US Government Printing Office, March 1983), chap. 5; and the articles on "The Persistent Problem of PVS" in *Hastings Center Report* 18 (February–March 1988): 26–47.
61. 61. Ramsey, *Ethics at the Edges of Life*, p. 172.
62. 62. President's Commission, *Deciding to Forego Life-Sustaining Treatment.*
63. 63. See John D. Lantos and Diane S. Lauderdale, *Preterm Babies, Fetal Patients, and Childbearing Choices* (Cambridge, MA: MIT Press, 2015), p. 150. For overviews of ethical issues in neonatal care, see Lantos, *The Lazarus Case: Life-and-Death Issues in Neonatal Care* (Baltimore, MD: Johns Hopkins University Press, 2001); Lantos and William L. Meadow, *Neonatal Bioethics: The Moral Challenges of Medical Innovation* (Baltimore, MD: Johns Hopkins University Press, 2006); Alan R. Fleischman, *Pediatric Ethics: Protecting the Interests of Children* (New York: Oxford University Press, 2016), chap. 4; and Dominic Wilkinson, *Death or Disability? The 'Carmentis Machine' and Decision-Making for Critically Ill Children* (Oxford: Oxford University Press, 2013).
64. 64. For a discussion of a version of this condition, see E. G. Yan et al., "Treatment Decision-making for Patients with the Herlitz Subtype of Junctional Epidermolysis Bullosa," *Journal of Perinatology* 27 (2007): 307–11. According to Julian Savulescu, this is the "best example" of a condition that renders a life "intolerable and not worth living." See Savulescu, "Is It in Charlie Gard's Best Interest to Die?" *Lancet* 389 (May 13, 2017): 1868–69. The Nuffield Council on Bioethics uses the concept of "intolerability" to describe situations where life-sustaining treatment would not be in the baby's "best interests" because of the burdens imposed by "irremediable suffering." *Critical Care Decisions in Fetal and Neonatal Medicine: Ethical Issues* (London: Nuffield Council on Bioethics, 2006).
65. 65. Lantos and Meadow, *Neonatal Bioethics*, pp. 16–17.
66. 66. Much of the support for the harm standard, as a replacement of, or as a supplement to, the best interest standard, builds on the work of Douglas S. Diekema, "Parental Refusals of Medical Treatment: The Harm Principle as Threshold for State Intervention," *Theoretical Medicine and Bioethics* 25, no. 4 (2004): 243–64; and Diekema, "Revisiting the Best Interest Standard: Uses and Misuses," *Journal of Clinical Ethics* 22, no. 2 (2011): 128–33. He argues, and we agree, that the harm standard functions primarily to warrant state intervention rather than to guide deliberations.
67. 67. For several defenses of the best interest standard, close to ours in many respects, see the following in the *American Journal of Bioethics* 18, no. 8 (2018), which is largely devoted to the best interest standard, the harm standard, and other competing approaches: Johan Christiaan Bester, "The Harm Principle Cannot Replace the Best Interest Standard: Problems with Using the Harm Principle for Medical Decision Making for Children," pp. 9–19; Loretta M. Kopelman, "Why the Best Interest Standard Is Not Self-Defeating, Too Individualistic, Unknowable, Vague or Subjective," pp. 34–37; Thaddeus Mason Pope, "The Best Interest Standard for Health Care Decision Making: Definition and Defense," pp. 36–38; Peta Coulson-Smith, Angela Fenwick, and Anneke Lucassen, "In Defense of Best Interests: When Parents and

Clinicians Disagree," pp. 67–69. Among the several defenses of the harm standard in this issue is D. Micah Hester, Kellie R. Lang, Nanibaa' A. Garrison, and Douglas S. Diekema, "Agreed: The Harm Principle Cannot Replace the Best Interest Standard … but the Best Interest Standard Cannot Replace the Harm Principle Either," pp. 38–41. See also the Diekema articles in the previous note.

68. 68. See Frank A. Chervenak and Laurence B. McCullough, "Nonaggressive Obstetric Management," *JAMA: Journal of the American Medical Association* 261 (June 16, 1989): 3439–40; and their article "The Fetus as Patient: Implications for Directive versus Nondirective Counseling for Fetal Benefit," *Fetal Diagnosis and Therapy* 6 (1991): 93–100.
69. 69. This case and the accompanying commentaries appear in Alexander A. Kon, Angira Patel, Steven Leuthner, and John D. Lantos, "Parental Refusal of Surgery in an Infant with Tricuspid Atresia," *Pediatrics* 138, no. 5 (2016): e20161730.
70. 70. See Kon's comments in Kon, Patel, Leuthner, and Lantos, "Parental Refusal of Surgery in an Infant with Trisucpid Atresia."
71. 71. See Patel's comments in Kon, Patel, Leuthner, and Lantos, "Parental Refusal of Surgery in an Infant with Tricuspid Atresia."
72. 72. For a review of this case, see John D. Lantos, "The Tragic Case of Charlie Gard," *JAMA Pediatrics* 171, no. 10 (2017): 935–36.
73. 73. Savalescu, "Is It in Charlie Gard's Best Interest to Die?" 1868–69.
74. 74. Dominic Wilkinson, "Beyond Resources: Denying Parental Requests for Futile Treatment," *Lancet* 389 (May 13, 2017): 1866–67. Wilkinson and Savulescu feature the Charlie Gard case in their coauthored book, *Ethics, Conflict and Medical Treatment for Children: From Disagreement to Dissensus* (London: Elsevier, 2018).
75. 75. This is the tack taken by Seema K. Shah, Abby R. Rosenberg, and Douglas S. Diekema, "Charlie Gard and the Limits of Best Interests," *JAMA Pediatrics* 171, no. 10 (October 2017): 937–38. However, the harm standard, which they defend in place of the best-interest standard, at least in matters of state intervention, cannot escape value judgments.
76. 76. See Jeff McMahan, "Killing, Letting Die, and Withdrawing Aid," *Ethics* 103 (1993): 250–79; James Rachels, "Killing, Letting Die, and the Value of Life," in his *Can Ethics Provide Answers? And Other Essays in Moral Philosophy* (Lanham, MD: Rowman & Littlefield, 1997), pp. 69–79; Tom L. Beauchamp, "When Hastened Death Is Neither Killing nor Letting-Die," in *Physician-Assisted Dying*, ed. Timothy E. Quill and Margaret P. Battin (Baltimore: Johns Hopkins University Press, 2004), pp. 118–29; Joachim Asscher, "The Moral Distinction between Killing and Letting Die in Medical Cases," *Bioethics* 22 (2008): 278–85; David Orentlicher, "The Alleged Distinction between Euthanasia and the Withdrawal of Life-Sustaining Treatment: Conceptually Incoherent and Impossible to Maintain," *University of Illinois Law Review* (1998): 837–59; and various articles in Steinbock and Norcross, eds., *Killing and Letting Die*, 2nd ed.
77. 77. Although the term *assisted suicide* is often used, we use it only when unavoidable. We prefer broader language, such as "physician-assisted dying" or "physician-arranged dying," not because of a desire to find euphemisms but because the broader language provides a more accurate description. Although the term *suicide* has the advantage of indicating that the one whose death is brought about authorizes or performs the final act, other conditions such as prescribing and transporting fatal substances may be as causally relevant as the "final act" itself. For related conceptual problems, see Franklin G. Miller, Robert D. Truog, and Dan W. Brock, "Moral Fictions and Medical Ethics," *Bioethics* 24 (2010): 453–60; and Helene Starks, Denise Dudzinski, and Nicole White (from original text written by Clarence H. Braddock III with Mark R. Tonelli), "Physician Aid-in-Dying," *Ethics in Medicine*, University of Washington School of Medicine (2013), available at https://depts.washington.edu/bioethx/topics/pad.html (accessed July 2, 2018).
78. 78. Howard Brody, "Messenger Case: Lessons and Reflections," *Ethics-in-Formation* 5 (1995): 8–9; Associated Press, "Father Acquitted in Death of His Premature Baby," *New York Times*, Archives 1995, available at https://www.nytimes.com/1995/02/03/us/father-acquitted-in-death-of-his-premature-baby.html (accessed July 3, 2018); and John Roberts, "Doctor Charged for Switching Off His Baby's Ventilator," *British Medical Journal* 309 (August 13, 1994): 430. Subsequent to this case similar cases have arisen in several countries.

79. 79. Cf. the diverse array of arguments and conclusions in James Rachels, "Active and Passive Euthanasia," *New England Journal of Medicine* 292 (January 9, 1975): 78–80; Miller, Truog, and Brock, "Moral Fictions and Medical Ethics"; Roy W. Perrett, "Killing, Letting Die and the Bare Difference Argument," *Bioethics* 10 (1996): 131–39; Dan W. Brock, "Voluntary Active Euthanasia," *Hastings Center Report* 22 (March–April 1992): 10–22; and Tom L. Beauchamp, "The Medical Ethics of Physician-assisted Suicide," *Journal of Medical Ethics* 15 (1999): 437–39 (editorial). Many, perhaps most, of the books opposed to the legalization of physician-assisted death operate from the premise that the act of physician-assisted death is wrong because of the inviolability of human life or intrinsic evil of aiming at death, etc. See, for example, Keown, *Euthanasia, Ethics and Public Policy*; Neal M. Gorsuch, *The Future of Assisted Suicide and Euthanasia* (Princeton, NJ: Princeton University Press, 2006); and Nigel Biggar, *Aiming to Kill: The Ethics of Euthanasia and Assisted Suicide* (Cleveland, OH: Pilgrim Press, 2004). By contrast, see Kevin Yuill, *Assisted Suicide: The Liberal, Humanist Case against Legalization* (Houndsmills, Basingstoke, Hampshire, UK: Palgrave Macmillan, 2013), which is particularly concerned about the "coercive implications" of the legalization of physician-assisted death. For a pro-con debate, see Emily Jackson and John Keown, *Debating Euthanasia* (Portland, OR: Hart, 2012).
80. 80. See Joseph J. Fins, *A Palliative Ethic of Care: Clinical Wisdom at Life's End* (Sudbury, MA: Jones & Bartlett, 2006); and Joanne Lynn et al., *Improving Care for the End of Life: A Sourcebook for Health Care Managers and Clinicians* (New York: Oxford University Press, 2007).
81. 81. Oregon Death with Dignity Act, Ore. Rev. Stat. § 127.800, available at https://www.oregon.gov/oha/PH/PROVIDERPARTNERRESOURCES/EVALUATIONRESEARCH/DEATHWITHDIGNITYACT/Pages/ors.aspx (accessed July 3, 2018). This act explicitly rejects the language of "physician-assisted suicide." It prefers the language of a right patients have to make a "request for medication to end one's life in a humane and dignified manner."
82. 82. See Lawrence O. Gostin, "Deciding Life and Death in the Courtroom: From Quinlan to Cruzan, Glucksberg, and Vacco—A Brief History and Analysis of Constitutional Protection of the 'Right to Die,'" *JAMA: Journal of the American Medical Association* 278 (November 12, 1997): 1523–28; and Yale Kamisar, "When Is There a Constitutional Right to Die? When Is There *No* Constitutional Right to Live?" *Georgia Law Review* 25 (1991): 1203–42.
83. 83. For discussions, see Douglas Walton, *Slippery Slope Arguments* (Oxford: Clarendon, 1992); Govert den Hartogh, "The Slippery Slope Argument," in *A Companion to Bioethics*, 2nd ed., ed. Helga Kuhse and Peter Singer (Malden, MA: Wiley-Blackwell, 2009), pp. 321–31; Christopher James Ryan, "Pulling Up the Runaway: The Effect of New Evidence on Euthanasia's Slippery Slope," *Journal of Medical Ethics* 24 (1998): 341–44; Bernard Williams, "Which Slopes Are Slippery?" in *Moral Dilemmas in Modern Medicine*, ed. Michael Lockwood (Oxford: Oxford University Press, 1985), pp. 126–37; James Rachels, *The End of Life: Euthanasia and Morality* (Oxford: Oxford University Press, 1986), chap. 10; and Penney Lewis, "The Empirical Slippery Slope from Voluntary to Non-Voluntary Euthanasia," *Journal of Law, Medicine & Ethics* 35 (March 1, 2007): 197–210.
84. 84. See Timothy E. Quill and Christine K. Cassel, "Nonabandonment: A Central Obligation for Physicians," in *Physician-Assisted Dying: The Case for Palliative Care and Patient Choice*, ed. Quill and Battin, chap. 2.
85. 85. See Franklin G. Miller, Howard Brody, and Timothy E. Quill, "Can Physician-Assisted Suicide Be Regulated Effectively?" *Journal of Law, Medicine & Ethics* 24 (1996): 225–32. Defenders of slippery-slope arguments in this context include John Keown, *Euthanasia, Ethics and Public Policy: An Argument Against Legislation* (Cambridge: Cambridge University Press, lst ed., 2002, 2nd ed., 2018), which contends that experience in countries that have legalized either physician-assisted dying or voluntary euthanasia show the effects of both "logical" and "empirical" slippery slopes; J. Pereira, "Legalizing Euthanasia or Assisted Suicide: The Illusion of Safeguards and Controls," *Current Oncology* 18 (April 2011): e38–45; and David Albert Jones, "Is There a Logical Slippery Slope from Voluntary to Nonvoluntary Euthanasia?" *Kennedy Institute of Ethics Journal* 21 (2011): 379–404; B. H. Lerner and A. L. Caplan, "Euthanasia in Belgium and the Netherlands: On a Slippery Slope?" *JAMA Internal Medicine* 175 (2015): 1640–41; William G. Kussmaul III, "The Slippery Slope of Legalization of Physician-Assisted Suicide," *Annals of Internal Medicine* 167, no. 8 (October 17, 2017): 595–96.

Critics of slippery-slope arguments include L. W. Sumner, *Assisted Death: A Study in Ethics and Law* (New York: Oxford University Press, 2011); Stephen W. Smith, "Fallacies of the Logical Slippery Slope in the Debate on Physician-Assisted Suicide and Euthanasia," *Medical Law Review* 13, no. 2 (July 1, 2005): 224–43; and Report of the Royal Society of Canada Expert Panel, *End-of-Life Decision Making* (Ottawa, ON: Royal Society of Canada, December 2011), available at http://rsc.ca/en/expert-panels/rsc-reports/end-life-decision-making (accessed July 4, 2018). After examining the laws and practical experience of jurisdictions around the world that authorize assisted dying in some cases, the latter concludes: "Despite the fears of opponents, it is … clear that the much-feared slippery slope has not emerged following decriminalization, at least not in those jurisdictions for which evidence is available" (p. 90).

86. 86. See, for example, Timothy E. Quill, "Legal Regulation of Physician-Assisted Death—The Latest Report Cards," *New England Journal of Medicine* 356 (May 10, 2007): 1911–13; Susan Okie, "Physician-Assisted Suicide—Oregon and Beyond," *New England Journal of Medicine* 352 (April 21, 2005): 1627–30; Courtney Campbell, "Ten Years of 'Death with Dignity,'" *New Atlantis* (Fall 2008): 33–46; and National Academies of Sciences, Engineering, and Medicine, *Physician-Assisted Death: Scanning the Landscape: Proceedings of a Workshop* (Washington, DC: National Academies Press, 2018).
87. 87. The information in this paragraph appears in the annual reports by the Oregon Health Authority. The Oregon Death with Dignity Act requires the Oregon Health Authority to publish information about patients and physicians who participate under the act, including the publication of an annual statistical report. See Oregon Health Authority, *Oregon Death with Dignity Act 2017 Data Summary*, as published in February 2018, available at https://www.oregon.gov/oha/PH/PROVIDERPARTNERRESOURCES/EVALUATIONRESEARCH/DEATHWITHDIGNITYACT/Documents/year20.pdf (accessed June 29, 2018). See also *The Oregon Death with Dignity Act: A Guidebook for Health Care Professionals Developed by the Task Force to Improve the Care of Terminally-Ill Oregonians*, convened by The Center for Ethics in Health Care, Oregon Health & Science University, 1st ed. (print), March 1998; current ed. (2008 online), available at http://www.ohsu.edu/xd/education/continuing-education/center-for-ethics/ethics-outreach/upload/Oregon-Death-with-Dignity-Act-Guidebook.pdf (accessed June 29, 2018). Many Oregonians are opposed to the Oregon law, but many others believe that it does not go far enough because it in effect excludes many persons with Alzheimer's, Parkinson's, Huntington's, multiple sclerosis and various other degenerative diseases, at least until their deaths are predicted to occur within six months.
88. 88. See Udo Schüklenk et al., "End-of-Life Decision-making in Canada: The Report by the Royal Society of Canada Expert Panel on End-of-life Decision-making," *Bioethics* 25 (2011) Suppl 1:1–73. This Expert Panel examines the international experience with laws authorizing assisted dying; Guenter Lewy, *Assisted Death in Europe and America: Four Regimes and Their Lessons* (New York: Oxford University Press, 2011); and the often updated information on various national policies at the UK site, My Death–My Decision, "Assisted Dying in Other Countries," available at https://www.mydeath-mydecision.org.uk/info/assisted-dying-in-other-countries/ (accessed July 3, 2018).
89. 89. See Bernard Gert, James L. Bernat, and R. Peter Mogielnicki, "Distinguishing between Patients' Refusals and Requests," *Hastings Center Report* 24 (July–August 1994): 13–15; Leigh C. Bishop et al., "Refusals Involving Requests" (Letters and Responses), *Hastings Center Report* 25 (July–August 1995): 4; Diane E. Meier et al., "On the Frequency of Requests for Physician Assisted Suicide in American Medicine," *New England Journal of Medicine* 338 (April 23, 1998): 1193–201; and Gerald Dworkin, Raymond G. Frey, and Sissela Bok, *Euthanasia and Physician-Assisted Suicide: For and Against* (New York: Cambridge University Press, 1998).
90. 90. As of July, 2018, physician-assisted death had been legalized in eight legal jurisdictions in the United States, whether through legislation, referendum, or a state supreme court decision: Oregon, Washington, Montana, Vermont, California, Colorado, District of Columbia, and Hawaii. For an overview, see Ezekiel J. Emanuel et al., "Attitudes and Practices of Euthanasia and Physician-Assisted Suicide in the United States, Canada, and Europe," *JAMA: Journal of the American Medical Association* 316, no. 1 (2016): 79–90. For another overview, from a variety of perspectives, see National Academies of Sciences, Engineering, and Medicine, *Physician-Assisted Death: Scanning the Landscape: Proceedings of a Workshop*.

91. 91. Cf. Allen Buchanan, "Intending Death: The Structure of the Problem and Proposed Solutions," in *Intending Death*, ed. Beauchamp, esp. pp. 34–38; Frances M. Kamm, "Physician-Assisted Suicide, the Doctrine of Double Effect, and the Ground of Value," *Ethics* 109 (1999): 586–605; and Matthew Hanser, "Why Are Killing and Letting Die Wrong?" *Philosophy and Public Affairs* 24 (1995): 175–201.
92. 92. Many moral arguments for justified physician-aid-in-dying focus on the relief of pain and suffering. However the "end of life concerns" most frequently listed by persons in Oregon who used their prescribed medication to end their lives were the following: diminished ability to engage in activities making life enjoyable (88.1%), loss of autonomy (87.4%), loss of dignity (67.1%), burden on family, friends, or caregivers (55.2%), and loss of control of bodily functions (37.1%). Only 21% listed inadequate pain control or concern about it. Oregon Health Authority, *Oregon Death with Dignity Act 2017 Data Summary.*
93. 93. *New York Times*, June 6, 1990, pp. A1, B6; June 7, 1990, pp. A1, D22; June 9, 1990, p. A6; June 12, 1990, p. C3; *Newsweek*, June 18, 1990, p. 46. Kevorkian's own description is in his *Prescription: Medicide* (Buffalo, NY: Prometheus Books, 1991), pp. 221–31. He was later convicted and served time in prison not for his more than one hundred acts of assisting in a person's suicide but for a single case of actively killing a patient (voluntary euthanasia). See Michael DeCesare, *Death on Demand: Jack Kevorkian and the Right-to-Die Movement* (Lanham, MD: Rowman & Littlefield, 2015).
94. 94. Timothy E. Quill, "Death and Dignity: A Case of Individualized Decision Making," *New England Journal of Medicine* 324 (March 7, 1991): 691–94, reprinted with additional analysis in Quill, *Death and Dignity* (New York: Norton, 1993); and Timothy Quill, *Caring for Patients at the End of Life: Facing an Uncertain Future Together* (Oxford: Oxford University Press, 2001).
95. 95. J. K. Kaufert and T. Koch, "Disability or End-of-Life: Competing Narratives in Bioethics," *Theoretical Medicine* 24 (2003): 459–69. See also Kristi L. Kirschner, Carol J. Gill, and Christine K. Cassel, "Physician-Assisted Death in the Context of Disability," in *Physician-Assisted Suicide*, ed. Robert F. Weir (Bloomington: Indiana University Press, 1997), pp. 155–66.
96. 96. For an examination of relevant US law, see Norman L. Cantor, *Making Medical Decisions for the Profoundly Mentally Disabled* (Cambridge, MA: MIT Press, 2005).
97. 97. See Hans-Martin Sass, Robert M. Veatch, and Rihito Kimura, eds., *Advance Directives and Surrogate Decision Making in Health Care: United States, Germany, and Japan* (Baltimore: Johns Hopkins University Press, 1998); Nancy M. P. King, *Making Sense of Advance Directives* (Dordrecht, Netherlands: Kluwer Academic, 1991; rev. ed. 1996); Peter Lack, Nikola Biller-Andorno, and Susanne Brauer, eds., *Advance Directives* (New York: Springer, 2014); and American Bar Association, "State Health Care Power of Attorney Statutes: Selected Characteristics January 2018," available at https://www.americanbar.org/content/dam/aba/administrative/law_aging/state-health-care-power-of-attorney-statutes.authcheckdam.pdf (accessed July 4, 2018).
98. 98. See, for example, the President's Council on Bioethics, *Taking Care: Ethical Caregiving in Our Aging Society* (Washington, DC: President's Council on Bioethics, 2005), chap. 2; Alasdair R. MacLean, "Advance Directives, Future Selves and Decision-Making," *Medical Law Review* 14 (2006): 291–320; A. Fagerlin and C. E. Schneider, "Enough: The Failure of the Living Will," *Hastings Center Report* 34, no. 2 (2004): 30–42; Dan W. Brock, "Advance Directives: What Is It Reasonable to Expect from Them?" *Journal of Clinical Ethics* 5 (1994): 57–60; Mark R. Tonelli, "Pulling the Plug on Living Wills: A Critical Analysis of Advance Directives," *Chest* 110 (1996): 816–22; David I. Shalowitz, Elizabeth Garrett-Mayer, and David Wendler, "The Accuracy of Surrogate Decision Makers: A Systematic Review," *Archives of Internal Medicine* 165 (2006): 493–97; Marcia Sokolowski, *Dementia and the Advance Directive: Lessons from the Bedside* (New York: Springer, 2018); and Lesley S. Castillo, Brie A. Williams, Sarah M. Hooper, et al., "Lost in Translation: The Unintended Consequences of Advance Directive Law on Clinical Care," *Annals of Internal Medicine* 154 (January 2011), available at http://annals.org/aim/article-abstract/746727/lost-translation-unintended-consequences-advance-directive-law-clinical-care (accessed July 4, 2018).
99. 99. See, for instance, Karen Detering and Maria J. Silveira (and Section Editor, Robert M. Arnold), "Advance Care Planning and Advance Directives," UpToDate (online), Wolters Kluwer, 2018, available at https://www.uptodate.com/contents/advance-care-planning-and-advance-directives (accessed July 4, 2018); Benjamin H. Levi and Michael J. Green, "Too Soon to Give Up: Re-Examining the Value of Advance Directives," *American Journal of Bioethics* 10 (April 2010): 3–22 (and responses thereafter);

Bernard Lo and Robert Steinbrook, "Resuscitating Advance Directives," *Archives of Internal Medicine* 164 (2004): 1501–6; Robert S. Olick, *Taking Advance Directives Seriously: Prospective Autonomy and Decisions near the End of Life* (Washington, DC: Georgetown University Press, 2001); and Joanne Lynn and N. E. Goldstein, "Advance Care Planning for Fatal Chronic Illness: Avoiding Commonplace Errors and Unwarranted Suffering," *Annals of Internal Medicine* 138 (2003): 812–18.

100. 100. See, for example, Joan M. Teno, Joanne Lynn, R. S. Phillips, et al., "Do Formal Advance Directives Affect Resuscitation Decisions and the Use of Resources for Seriously Ill Patients?" SUPPORT Investigators: Study to Understand Prognoses and Preferences for Outcomes and Risks of Treatments, *Journal of Clinical Ethics* 5 (1994): 23–30.
101. 101. Maria J. Silveira, Scott Y. H. Kim, and Kenneth M. Langa, "Advance Directives and Outcomes of Surrogate Decision Making before Death," *New England Journal of Medicine* 362 (April 1, 2010): 1211–18; Joan Teno, Joanne Lynn, Neil Wenger, et al., "Advance Directives for Seriously Ill Hospitalized Patients: Effectiveness with the Patient Self-Determination Act and the SUPPORT Intervention," *Journal of the American Geriatrics Society*, published April 2015, available at "https://onlinelibrary.wiley.com/doi/abs/10.1111/j.1532-5415.1997.tb05178.x (accessed July 4, 2018); and Karen M. Detering, Andrew D. Hancock, Michael C. Reade, and William Silvester, "The Impact of Advance Care Planning on End of Life Care in Elderly Patients: Randomised Controlled Trial," *BMJ: British Medical Journal* 340 (2010): c1345, available at https://www.ncbi.nlm.nih.gov/pmc/articles/PMC2844949/ (accessed June 30, 2018). Debate continues about whether advance directives have—or for that matter, should have—an impact on health care costs. See Douglas B. White and Robert M. Arnold, "The Evolution of Advance Directives," *JAMA: Journal of the American Medical Association* 306 (October 5, 2011): 1485–86.
102. 102. Su Hyun Kim and Diane Kjervik, "Deferred Decision Making: Patients' Reliance on Family and Physicians for CPR Decisions in Critical Care," *Nursing Ethics* 12 (2005): 493–506. For a fuller examination of the family in bioethical matters, see Hilde Lindemann Nelson and James Lindemann Nelson, *The Patient in the Family: An Ethic of Medicine and Families*, Reflective Bioethics (New York: Routledge, 1995).
103. 103. See Judith Areen, "The Legal Status of Consent Obtained from Families of Adult Patients to Withhold or Withdraw Treatment," *JAMA: Journal of the American Medical Association* 258 (July 10, 1987): 229–35; Charles B. Sabatino, "The Evolution of Health Care Advance Planning Law and Policy," *Milbank Quarterly* 88 (2010): 211–38; and American Bar Association, Commission on Law and Aging, "Health Care Decision Making"; see the relevant publications on surrogate decision making, available at https://www.americanbar.org/groups/law_aging/resources/health_care_decision_making.html (accessed July 4, 2018).
104. 104. Patricia King, "The Authority of Families to Make Medical Decisions for Incompetent Patients after the Cruzan Decision," *Law, Medicine & Health Care* 19 (1991): 76–79.
105. 105. Mark P. Aulisio, "Standards for Ethical Decision Making at the End of Life," in *Advance Directives and Surrogate Decision Making in Illinois*, ed. Thomas May and Paul Tudico (Springfield, IL: Human Services Press, 1999), pp. 25–26.
106. 106. For some significant subtleties, see Susan P. Shapiro, "Conflict of Interest at the Bedside," in *Conflict of Interest in Global, Public and Corporate Governance*, ed. Anne Peters and Lukas Handschin (Cambridge: Cambridge University Press, 2012), pp. 334–54.
107. 107. David Wendler, "The Effect on Surrogates of Making Treatment Decisions for Others," *Annals of Internal Medicine* 154 (March 1, 2011): 336–46.
108. 108. David E. Weissman, "Decision Making at a Time of Crisis Near the End of Life," *JAMA: Journal of the American Medical Association* 292 (2004): 1738–43.
109. 109. For an analysis of the role of courts and the connection to valid consent, see M. Strätling, V. E. Scharf, and P. Schmucker, "Mental Competence and Surrogate Decision-Making towards the End of Life," *Medicine, Health Care and Philosophy* 7 (2004): 209–15.
110. 110. See our discussions in Chapter 3, pp. 89–92, and Chapter 7, pp. 286–89.
111. 111. The facts of the case and observations about it are found in Peter Pronovost, Dale Needham, Sean Berenholtz, et al., "An Intervention to Decrease Catheter-Related Bloodstream Infections in the ICU," *New England Journal of Medicine* 355 (2006): 2725–32; and Mary Ann Baily, "Harming through Protection?" *New England Journal of Medicine* 358 (2008): 768–69.

112. 112. US Department of Health and Human Services, Office for Human Research Protections, *OHRP Statement Regarding the New York Times Op-Ed Entitled* "*A Lifesaving Checklist*," News, January 15, 2008, available at http://www.hhs.gov/ohrp/news/recentnews.html#20080215 (accessed December 5, 2011).
113. 113. See Holly Fernandez Lynch, Barbara E. Bierer, I. Glenn Cohen, and Suzanne M. Rivera, eds., *Specimen Science: Ethics and Policy Implications*, Basic Bioethics (Cambridge, MA: MIT Press, 2017).
114. 114. Allen Buchanan, "An Ethical Framework for Biological Samples Policy," in National Bioethics Advisory Commission, *Research Involving Human Biological Materials: Ethical Issues and Policy Guidance*, vol. 2 (Rockville, MD: National Bioethics Advisory Commission, January 2000); Christine Grady et al., "Broad Consent for Research with Biological Samples: Workshop Conclusions," *American Journal of Bioethics* 15 (2015): 34–42; Teddy D. Warner et al., "Broad Consent for Research on Biospecimens: The Views of Actual Donors at Four U.S. Medical Centers," *Journal of Empirical Research on Human Research Ethics* (February 2018), available at http://journals.sagepub.com/doi/abs/10.1177/1556264617751204 (accessed July 5, 2018); Karen J. Maschke, "Wanted: Human Biospecimens," *Hastings Center Report* 40, no. 5 (2010): 21–23; and Rebecca D. Pentz, Laurent Billot, and David Wendler, "Research on Stored Biological Samples: Views of African American and White American Cancer Patients," *American Journal of Medical Genetics*, published online March 7, 2006, http://onlinelibrary.wiley.com/doi/10.1002/ajmg.a.31154/full. For a thorough examination, from various perspectives, of broad consent as well as other ethical issues in research on biospecimens, such as privacy, justice, and governance, see Lynch, Bierer, Cohen, and Rivera, eds., *Specimen Science: Ethics and Policy Implications*, which includes an adapted version of Grady et al., "Broad Consent for Research with Biological Samples," pp. 167–84.
115. 115. *Havasupai Tribe of Havasupai Reservation v. Arizona Bd. of Regents*, 204 P.3d 1063 (Ariz. Ct. App. 2008); Dan Vorhaus, "The Havasupai Indians and the Challenge of Informed Consent for Genomic Research," *The Privacy Report*, available at http://www.genomicslawreport.com/index.php/2010/04/21/the-havasupai-indians-and-the-challenge-of-informed-consent-for-genomic-research/ (accessed June 30, 2018); Amy Harmon, "Indian Tribe Wins Fight to Limit Research of Its DNA," *New York Times*, April 21, 2010, p. A1, available at http://www.nytimes.com/2010/04/22/us/22dna.html (accessed June 30, 2018); and Amy Harmon, "Havasupai Case Highlights Risks in DNA Research," *New York Times*, April 22, 2010, available at http://www.nytimes.com/2010/04/22/us/22dnaside.html (accessed June 30, 2018).
116. 116. See Michelle M. Mello and Leslie E. Wolf, "The Havasupai Indian Tribe Case—Lessons for Research Involving Stored Biological Samples," *New England Journal of Medicine* 363 (July 15, 2010): 204–7; American Indian and Alaska Native Genetics Resources, National Congress of American Indians, "Havasupai Tribe and the Lawsuit Settlement Aftermath," available at http://genetics.ncai.org/case-study/havasupai-Tribe.cfm (accessed July 4, 2018); and Nanibaa' A. Garrison and Mildred K. Cho, "Awareness and Acceptable Practices: IRB and Researcher Reflections on the Havasupai Lawsuit," *AJOB Primary Research* 4 (2013): 55–63.
117. 117. Amy Harmon, "Where'd You Go with My DNA?" *New York Times*, April 25, 2010, available at http://www.nytimes.com/2010/04/25/weekinreview/25harmon.html?ref=us (accessed June 30, 2018).

6

Beneficence

We have seen in the last two chapters that morality requires that we treat persons autonomously and refrain from harming them, but morality also requires that we contribute to their welfare. Principles of beneficence potentially demand more than the principle of nonmaleficence, because agents must take positive steps to help others, not merely refrain from harmful acts. An implicit assumption of beneficence undergirds all medical and health care professions and their institutional settings. For example, attending to the welfare of patients—not merely avoiding harm—is at the heart of medicine's goal, rationale, and justification. Likewise, preventive medicine, public health, and biomedical research embrace values of public beneficence.

We examine *two principles of beneficence* in this chapter: positive beneficence and utility. The principle of positive beneficence requires agents to provide benefits to others. The principle of utility requires agents to balance benefits, risks, and costs to produce the best overall results. We also explore the virtue of benevolence, obligatory beneficence, and nonobligatory ideals of beneficence. We then show how to handle conflicts between beneficence and respect for autonomy that occur in paternalistic refusals to accept a patient's wishes and in public policies designed to protect or improve individuals' health. Thereafter, this chapter focuses on proposals to balance benefits, risks, and costs through analytical methods designed to implement the principle of utility in both health policy and clinical care. We conclude that these analytical methods have a useful, albeit limited, role as aids in decision making.

The Concept of Beneficence and Principles of Beneficence

In ordinary English, the term *beneficence* connotes acts or qualities of mercy, kindness, friendship, generosity, charity, and the like. We use the term in this chapter to cover beneficent action in a broad sense to include all norms, dispositions, and actions with the goal of benefiting or promoting the well-being of other persons. *Benevolence* refers to the character trait, or virtue, of being disposed to act for the benefit of others. *Principle of beneficence* refers to a statement of a general moral obligation to act for the benefit of others. Many morally commendable acts of beneficence are not obligatory, but some are obligatory.

Beneficence and benevolence have played central roles in certain ethical theories. Utilitarianism, for example, is built on the single principle of beneficence referred to as the principle of utility. During the Scottish Enlightenment, major figures, including Francis Hutcheson and David Hume, made benevolence the centerpiece of their common-morality theories. Some of these theories closely associate benefiting others with the goal of morality itself. We concur that obligations to confer benefits, to prevent and remove harms, and to weigh an action's possible goods against its costs and possible harms are central to the moral life. However, principles of beneficence are not sufficiently foundational to ground *all other* moral principles and rules in the way many utilitarians have maintained. (See further our discussion of utilitarian theory in Chapter 9, pp. 389–94.)

The principle of utility in our account in this chapter is therefore not identical to the classic utilitarian principle of utility. Whereas utilitarians view utility as a fundamental, absolute principle of ethics, we treat it as one among a number of equally important prima facie principles. The principle of utility that we defend is legitimately overridden by other moral principles in a variety of circumstances, and likewise it can override other prima facie principles under various conditions.

Obligatory Beneficence and Ideal Beneficence

Some deny that morality imposes positive obligations of beneficence. They hold that beneficence is purely a virtuous ideal or an act of charity, and thus that persons do not violate obligations of beneficence if they fail to

act beneficently.[1] These views rightly indicate a need to clarify insofar as possible the points at which beneficence is optional and the points at which it is obligatory.

An instructive and classic example of this problem appears in the New Testament parable of the Good Samaritan, which illustrates several problems in interpreting beneficence. In this parable, robbers beat and abandon a "half-dead" man traveling from Jerusalem to Jericho. After two travelers pass by the injured man without rendering help, a Samaritan sees him, feels compassion, binds up his wounds, and brings him to an inn to take care of him. In having compassion and showing mercy, the Good Samaritan expressed an attitude of caring about the injured man and he also took care of him. Both the Samaritan's motives and his actions were beneficent. Common interpretations of the parable suggest that positive beneficence is here an ideal rather than an obligation, because the Samaritan's act seems to exceed ordinary morality. But even if the case of the Samaritan does present an ideal of conduct, there are some obligations of beneficence.

Virtually everyone agrees that the common morality does not contain a principle of beneficence that requires severe sacrifice and extreme altruism—for example, putting one's life in grave danger to provide medical care or donating both of one's kidneys for transplantation. Only ideals of beneficence incorporate such extreme generosity. Likewise, we are not morally required to benefit persons on all occasions, even if we are in a position to do so. For example, we are not morally required to perform all possible acts of generosity or charity that would benefit others. Much beneficent conduct constitutes ideal, rather than obligatory, action; and the line between an obligation of beneficence and a moral ideal of beneficence is often unclear. (See our treatment of this subject in the section on supererogation in Chapter 2, pp. 46–48.)

The principle of positive beneficence supports an array of prima facie rules of obligation, including the following:

1. 1. Protect and defend the rights of others.
2. 2. Prevent harm from occurring to others.
3. 3. Remove conditions that will cause harm to others.
4. 4. Help persons with disabilities.
5. 5. Rescue persons in danger.

Distinguishing Rules of Beneficence from Rules of Nonmaleficence

Rules of beneficence differ in several ways from those of nonmaleficence. In Chapter 5 we argued that rules of nonmaleficence (1) are negative prohibitions of action, (2) must be followed impartially, and (3) provide moral reasons for legal prohibitions of certain forms of conduct. By contrast, rules of beneficence (1) present positive requirements of action, (2) need not always be followed impartially, and (3) generally do not provide reasons for legal punishment when agents fail to abide by them.

The second condition of impartial adherence asserts that we are morally prohibited by rules of nonmaleficence from causing harm to anyone. We are obligated to act nonmaleficently toward all persons at all times (although the principle of nonmaleficence is sometimes justifiably overridden when it comes into conflict with other principles). By contrast, obligations of beneficence often permit us to help or benefit those with whom we have special relationships when we are not required to help or benefit those with whom we have no such relationship. With family, friends, and others of our choice, morality ordinarily allows us to practice beneficence with partiality. Nonetheless, we will show that we are obligated to follow impartially *some* rules of beneficence, such as those requiring efforts to rescue strangers when rescue efforts pose little risk to the prospective rescuer.

General and Specific Beneficence

A distinction between specific and general beneficence dispels some of the confusion surrounding the distinction between obligatory beneficence and nonobligatory moral ideals of beneficence. Specific beneficence usually rests on moral relations, contracts, or special commitments and is directed at particular parties, such as children, friends, contractors, or patients. For instance, many specific obligations of beneficence in health care—often

referred to as duties—rest on a health professional's assumption of obligations through entering a profession and taking on professional roles. By contrast, general beneficence is directed beyond special relationships to all persons.

Virtually everyone agrees that all persons are obligated to act in the interests of their children, friends, and other parties in special relationships. The role responsibilities of health professionals to take care of patients and subjects provide many examples. However, the idea of a *general* obligation of beneficence is more controversial. W. D. Ross suggests that obligations of general beneficence "rest on the mere fact that there are other beings in the world whose condition we can make better."[2] From this perspective, general beneficence obligates us to benefit persons whom we do not know or with whose views we are not sympathetic. The notion that we have the same impartial obligations of beneficence to innumerable persons we do not know as we have to our families is overly demanding and impractical. It is also perilous because this standard may divert attention from our obligations to those with whom we have special moral relationships, and to whom our responsibilities are clear rather than indefinite. The more widely we generalize obligations of beneficence, the less likely we will be to meet our primary responsibilities. For this reason, among others, the common morality recognizes significant limits to the scope of general obligations of beneficence.

Some writers try to set these limits by distinguishing between the removal of harm, the prevention of harm, and the promotion of benefit. In developing a principle of "the obligation to assist," Peter Singer has throughout his career been interested in how to reduce the evils of global harm and suffering in the most effective manner. He distinguishes preventing evil from promoting good, and contends that "if it is in our power to prevent something bad from happening, without thereby sacrificing anything of comparable moral importance, we ought, morally, to do it."[3] His major argument is that serious shortages of food, shelter, and health care threaten human life and welfare but are preventable. If any given person has some capacity to act to prevent these evils—for example, by donation to aid agencies—without loss of goods of comparable importance, this person acts unethically by not contributing to alleviate these shortages. Singer's key point is that in the face of preventable disease and poverty we are morally obligated to donate time or resources toward their eradication until we reach a level at which, by giving more, we would cause as much suffering to ourselves as we would relieve through our gift. This highly demanding principle of beneficence requires all of us with the power to do so to invest in rescuing needy persons globally.

Singer's criterion of comparable importance sets a limit on sacrifice: We ought to donate time and resources until we reach a level at which, by giving more, we would sacrifice something of comparable moral importance. At this level of sacrifice we might cause as much suffering to ourselves as we would relieve through our gift. While Singer leaves open the question of what counts as comparably morally important, his argument implies that morality sometimes requires us to make large personal sacrifices to rescue needy persons around the world. As judged by common-morality standards, this account is overdemanding, even though it sets forth an admirable moral ideal. The requirement that persons seriously disrupt reasonable life plans in order to benefit the sick, undereducated, or starving exceeds the limits of basic obligations. In short, Singer's principle expresses a commendable moral ideal of beneficence, but it is doubtful that the principle can be justifiably claimed to be a general obligation of beneficence.

Singer resists this assessment. He regards ordinary morality as endorsing a demanding harm prevention principle. He assesses the almost universal lack of a commitment to contribute to poverty relief as a failure to draw the correct implications from the moral principle(s) of beneficence that all moral persons accept. We respond, constructively, to this line of argument in the next section, where we treat the limits of obligations of rescue. The claim that Singer-type beneficence makes excessively strong demands is, we will argue, best tested by these rescue cases. We offer a five-condition analysis of beneficence that we judge more satisfactory than Singer's principle.

Singer has countered objections that his principle sets an overly demanding standard. Although he still adheres to his exacting principle of beneficence, he acknowledges that it may be maximally productive to *publicly advocate* a less demanding principle. He has suggested a percentage of income such as 10%, which is more than a small donation, but not so large as to be at the elevated level of a saint.[4] This revised thesis more appropriately

sets limits on the scope of the obligation of beneficence—limits that reduce required costs and impacts on the agent's life plans and that make meeting one's obligations a realistic possibility.

Singer also has offered more complicated formulas about how much one should donate and has sought to identify the social conditions that motivate people to give.[5] He responds to critics[6] by conceding that the limit of what we should publicly advocate as a level of giving is a person's "fair share" of what is needed to relieve poverty and other problems. A fair share may be more or may be less than his earlier formulations suggested, but Singer seems to view the fair-share conception as a realistic goal. His attention to motivation to contribute to others illuminates one dimension of the nature and limits of beneficence. Of course, obligation and motivation are distinguishable, and, as Singer appreciates, it will prove difficult in many circumstances to motivate people to live up to their obligations (as Singer conceives them) to rescue individuals in need.

The Duty of Rescue as Obligatory Beneficence

Some circumstances eliminate discretionary choice regarding beneficiaries of our beneficence. Consider the stock example of a passerby who observes someone drowning but stands in no special moral relationship to the drowning person. The obligation of beneficence is not sufficiently robust to require a passerby who is a poor swimmer to risk his or her life by trying to swim a hundred yards to rescue someone drowning in deep water. Nonetheless, the passerby who is well-placed to help the victim in some way, without incurring a significant risk to himself or herself, has a moral obligation to do so. If the passerby does nothing—for example, fails to alert a nearby lifeguard or fails to call out for help—the failure is morally culpable as a failure of obligation. The obligation to help here, in the absence of significant risk or cost to the agent, eliminates the agent's discretionary choice.

Apart from close moral relationships, such as contracts or the ties of family or friendship, we propose that a person X has a prima facie obligation of beneficence, in the form of a duty of rescue, toward a person Y if and only if each of the following conditions is satisfied (assuming that X is aware of the relevant facts):[7]

1. 1. Y is at risk of significant loss of or damage to life, health, or some other basic interest.
2. 2. X's action is necessary (singly or in concert with others) to prevent this loss or damage.
3. 3. X's action (singly or in concert with others) will probably prevent this loss or damage.[8]
4. 4. X's action would not present significant risks, costs, or burdens to X.
5. 5. The benefit that Y can be expected to gain outweighs any harms, costs, or burdens that X is likely to incur.

Although it is difficult to state the precise meaning of "significant risks, costs, or burdens" in the fourth condition, reasonable thresholds can be set, and this condition, like the other four, is essential to render the action *obligatory* on grounds of beneficence (by contrast to a nonobligatory act of beneficence).

We can now investigate the merit of these five conditions of obligatory beneficence by using three test cases. The first is a borderline case of specific obligatory beneficence, involving rescue, whereas the second presents a clear-cut case of specific obligatory beneficence. The third, a hypothetical case, directs our attention to obligations of beneficence when it is possible to help only some members of a group at risk in an epidemic. After addressing these cases, we consider the possibility of a duty to rescue in the context of research.

In the first case, which we introduced in Chapter 5 (pp. 157–58), Robert McFall was diagnosed as having aplastic anemia, which is often fatal, but his physician believed that a bone marrow transplant from a genetically compatible donor could increase his chances of surviving. David Shimp, McFall's cousin, was the only relative willing to undergo the first test, which established tissue compatibility. Shimp then unexpectedly refused to undergo the second test for genetic compatibility. When McFall sued to force his cousin to undergo the second test and to donate bone marrow if he turned out to be compatible, the judge ruled that the *law* did not allow him to force Shimp to engage in such acts of positive beneficence. However, the judge also stated his view that Shimp's refusal was "*morally* indefensible."

The judge's moral assessment is questionable because it is unclear that Shimp shirked an obligation. Conditions 1 and 2 listed previously were met for an obligation of specific beneficence in this case, but condition 3 was not satisfied. McFall's chance of surviving one year (at the time) would have increased only from 25% to between 40% and 60%. These contingencies make it difficult to determine whether a principle of beneficence can be validly specified so that it demands a particular course of action in this case. Although most medical commentators agreed that the risks to the donor were minimal in this case, Shimp was concerned about condition 4. Bone marrow transplants, he was told, require 100 to 150 punctures of the pelvic bone. These punctures can be painlessly performed under anesthesia, and the major risk at the time was a 1 in 10,000 chance of death from anesthesia. Shimp, however, believed that the risks were greater ("What if I become a cripple?" he asked) and that they outweighed the probability and magnitude of benefit to McFall. This case, all things considered, seems to be a borderline case of obligatory beneficence.

In the *Tarasoff* case, discussed in Chapter 1 (pp. 10–11), a therapist, on learning of his patient's intention to kill an identified woman, notified the police but did not warn the intended victim because of constraints of confidentiality. Suppose we modify the actual circumstances in the *Tarasoff* case to create the following hypothetical situation: A psychiatrist informs all of his patients that he may not keep information confidential if serious threats to other persons are disclosed by the patient. The patient agrees to treatment under these conditions and subsequently reveals an unmistakable intention to kill an identified woman. The psychiatrist may now either remain aloof and maintain confidentiality or take measures to protect the woman by notifying her or the police, or both. What does morality—and specifically beneficence—demand of the psychiatrist in this case?

Only a remarkably narrow account of moral obligation would assert that the psychiatrist is under no obligation to protect the woman by contacting her or the police or both. The psychiatrist is not at significant risk and will suffer virtually no inconvenience or interference with his life plans. If morality does not demand this much beneficence, it is difficult to see how morality imposes any positive obligations at all. Even if a competing obligation exists, such as protection of confidentiality, requirements of beneficence will, in the hypothetical case we have constructed, override the obligation of confidentiality. In similar situations, health care professionals may have an overriding moral obligation to warn spouses or lovers of HIV-infected patients who refuse to disclose their status to their partners and who refuse to engage in safer sex practices.

What is the morally relevant difference between these rescue cases involving individuals and those discussed in the previous section? We there suggested that rescuing a drowning person involves a specific obligation not present with global poverty, because the rescuer is "well-placed at that moment to help the victim." However, many of us are well placed to help people in poverty by giving modest sums of money. We can do so at little risk to ourselves and with some chance of limited benefit to others. One response is that in the drowning case there is a specific individual toward whom we have an obligation, whereas in the poverty cases we have obligations toward entire populations of people, only a few of whom we can possibly hope to help through a gift.

It is tempting to suppose that we are obligated to act only when we can help specific, identifiable individuals, not when we can help only some of the members of a larger group. However, this line of argument has implausible implications, particularly when the size of groups is smaller in scale. Consider a situation in which an epidemic breaks out in a reasonably small community, calling for immediate quarantine, and hundreds of persons who are not infected cannot return to their homes if infected persons are in the home. They are also not allowed to leave the city limits, and all hotel rooms are filled. Authorities project that you could prevent the deaths of approximately twenty noninfected persons by offering them portable beds (supplied by the city) in your house. Conditions would become unsanitary if more than twenty persons were housed in one home, but there are enough homes to house every stranded person if each house in the community takes twenty persons. It seems implausible to say that no person is morally obligated to open his or her house to these people for the weeks needed to control the epidemic, even though no one person has a specific obligation to any one of the stranded individuals. The hypothesis might be offered that this obligation arises only because they are all members of the community, but this hypothesis is implausible because it would arbitrarily exclude visitors who were stranded.

It is doubtful that ethical theory and practical deliberation can establish precise, determinate limits on the scope of obligations of beneficence. Attempts to do so will involve setting a revisionary line in the sense that they will draw a sharper boundary for our obligations than the common morality recognizes. Although the limits of

beneficence are certainly not precise, we have argued in this section that we can still appropriately fix or specify obligations of beneficence in some situations.

We will now connect these conclusions about the duty to rescue to a difficult ethical problem in policies and programs in research.

Expanded Access and Continued Access in Research

An excellent test for our analysis of obligations of beneficence and the duty of rescue is found in programs and policies of expanded access and continued access to investigational (experimental) products such as drugs and medical devices.

Expanded access to investigational products. In the absence of effective ways to treat serious medical conditions, many patients and their families are keenly interested in gaining access to promising drugs or devices that are in clinical trials but have not yet been approved. Societal perceptions of clinical research have shifted significantly over the last few decades. Beginning in the 1980s, especially as a result of the efforts of AIDs activists, increasing access to clinical trials became a major goal.[9] But not everyone with a particular medical condition meets the criteria for eligibility to participate in a clinical trial on treatments for their condition. In the United States, the Food and Drug Administration (FDA) undertook several initiatives to expedite the process of making new drugs available to treat serious conditions that lack effective alternative treatments. These initiatives use designations such as "fast track," "breakthrough therapy," "accelerated approval," and "priority review."[10]

The main moral question is whether it is sometimes either morally acceptable or morally obligatory to provide an investigational product to seriously ill patients such as persons with life-threatening conditions who cannot enroll in a clinical trial and who cannot wait until a promising product receives approval. Policies that do so are commonly called either "expanded access" or "compassionate use" programs. The two terms are not synonymous, but they both identify the same type of program, namely, one that authorizes access to an investigational product that does not yet have regulatory approval but that has passed basic safety tests (Phase I) and remains within the approval process.[11]

In a positive response to complaints that its program of "expanded access" is too cumbersome and too slow, the FDA has streamlined its procedures for application and access. Complaints and related concerns led to the passage of a federal "right to try" law in 2018 (similar to several state laws).[12] This legislation, which provides an option beyond a clinical trial or the FDA's explicit "expanded access" program, is expected to increase the number of terminally ill patients who are able to access investigational treatments. However, critics charge that this legislation often creates false hopes and threatens to delay or undermine the process of clinical research needed to determine both the safety and efficacy of investigational treatments. Some critics also charge that this legislation is part of a broader effort to subvert government regulation of the pharmaceutical industry.[13]

The primary goal of clinical research is scientific understanding that can lead to sound clinical interventions. Research is generally aimed at ensuring that potential treatments are safe and efficacious, not at immediately providing treatments. Research on new products therefore does not carry clinical obligations of health care, and clinical investigators and research sponsors are not morally obligated to provide access to an investigational product outside of a clinical trial. Sometimes, however, the following circumstances occur: A program of expanded access, based on available data, is reasonably safe and could possibly benefit some patients; no alternative therapies are available; and therapeutic use of the product does not threaten the scheduled completion or the results of a clinical trial. In these cases, it is morally permissible to adopt a program of expanded access, and in some cases investigational treatments have worked for patients enrolled in these programs. The use of the drug AZT in the treatment of AIDS is a classic case in which compassionate use would have been justified had there been an adequate supply of the drug available at the time. (See our discussion of this case in Chapter 8, pp. 366–67.)

Part of the reason for the virtue-grounded language of "compassionate use" is that though it is clearly compassionate and justified to provide some investigational products for therapeutic use, it is generally not obligatory to do so. In some cases, it is even obligatory not to provide access either because the risks are too high for patients or because access might seriously endanger clinical trial goals. Most investigational products do not survive clinical trials to achieve regulatory approval, and many turn out to have harmful side effects. If it is justified to proceed with a "compassionate use" program, the justification will likely appeal to a moral ideal, as analyzed in Chapter 2, rather than a moral obligation. It would be obligatory to undertake an expanded access program only if the situation conformed to all five conditions in the analysis of a duty of rescue that we discussed in the previous section.

In the normal course of investigational products, the prospect that all five conditions will be satisfied in any given new case is unlikely. In most potential compassionate use programs, condition 3 (will probably prevent a loss), condition 4 (will not present significant risks, costs, or burdens), or condition 5 (potential benefit can be expected to outweigh harms, costs, or burdens likely to be incurred) will not be satisfied. Often predictions and hopes about innovative treatments are not met. An apt illustration comes from the experimental treatment of breast cancer with high-dose chemotherapy followed by bone marrow transplantation. Perceptible initial improvement using aggressive applications in early-phase trials led to requests for expanded access from many patients. Approximately 40,000 women were given expanded access to this investigational approach—despite weak evidence of efficacy—and only 1,000 women participated in the independent clinical trial. The completed clinical trial established that this investigational strategy provided no benefits over standard therapies and actually elevated the risk of mortality. In short, this expanded access program increased risks for thousands of patients without additional benefits.[14]

Condition 3 can involve notably complicated decision making. However, we can easily imagine an extraordinary circumstance, such as a public health emergency, in which all of these conditions are satisfied and create an ethical obligation, not merely a moral ideal, of rescue through expanded access. The unusual case of the antiviral drug ganciclovir represents an interesting clinical situation of compassionate use because it satisfies all five conditions of the duty of rescue independently of a clinical trial and yet only questionably created an obligation on the part of the pharmaceutical company to provide the product. Ganciclovir had been shown to work in the laboratory against a previously untreatable viral infection, but a clinical trial was still years away. Authorization was given for first use of the drug in a few emergency compassionate use cases. The drug was demonstrated to be efficacious by *evidence of a different nature* than the information collected in a clinical trial. For example, retinal photographs showed changes in eye infections after treatment.[15] Although the provision of ganciclovir in this compassionate use program was controversial from the beginning, the program in retrospect clearly was justified, even though it cannot be said to have been morally obligatory when initiated. Syntex, the pharmaceutical company that developed the drug, created what would become a five-year expanded access program. The company was trapped into continuing the program, which it had planned to be only short term, because the US FDA would not approve ganciclovir in the absence of a scientific trial.

In sum, expanding patients' access to investigational products can sometimes be permissible beneficence, and it can occasionally be obligatory beneficence (when our listed conditions are met). By contrast, *continued access* to investigational products, a related but notably different practice, is more likely to be an obligation of specific beneficence, as we will now see.

Continued access to investigational products. The moral problem of continued access is how to identify the conditions under which it is morally obligatory, after a clinical trial has ended, to continue to provide an investigational product to research subjects who favorably responded to the product during the trial. Continued access may occur in several ways. The former subjects in the trial might continue as subjects in an extension of the trial on the same product or they might simply be given the product by the research sponsor. When subjects have responded favorably to an investigational product during the course of a trial and their welfare interests will be set back if the effective intervention is no longer available to them, two moral considerations distinguish this situation from that of expanded access. First, our analysis of the principle of nonmaleficence in Chapter 5 suggests that sponsors and investigators would be causing harm to research subjects by denying them further access to a product that is helping them address serious health problems or avoid death. Second, obligations of

reciprocity (a moral notion treated in the next section in the present chapter) suggest that research subjects are owed access to an apparently successful treatment at the end of their service in a clinical trial because they undertook risks to help produce knowledge about that product to help patients, which is also knowledge that advances science and benefits sponsors and investigators involved in the research.

These two moral considerations differentiate continued access from expanded access. They warrant the conclusion that there can be—and we think frequently are—moral obligations to provide continued access to investigational products for former research subjects. These obligations are independent of those created by our five-condition analysis of the duty of rescue. Although most of these five conditions are satisfied in many cases of continued access, condition 3 (will probably prevent loss or damage) often is not satisfied. Our view is that even if condition 3 is not satisfied, there still can be sufficient moral grounds to create an obligation to provide a continued access program because of demands of reciprocity and nonmaleficence. These moral grounds apply when there is good evidence that the research subject is currently benefiting even if there is inconclusive evidence that he or she will benefit in the long run.

Unlike the ordinary expanded access situation, it is unethical to withdraw an effective investigational product from a research subject who has a serious disorder or faces a significant risk of death and who has responded favorably to the investigational product. Sponsors and investigators should make conscientious efforts before a trial begins to ensure that a program of continued access is in place for all subjects for whom an investigational product proves effective. They also have obligations to state the conditions of continued access in the research protocol and to inform all potential subjects as part of the consent process what will happen if they respond favorably to the investigational products. Disclosures should be made regarding both the nature and duration of the continued access program, as well as the source of financing. If a protocol and consent form lack such information, the review committee should require investigators to justify the omission.[16]

However, these conclusions need a proviso. In some cases, a product under study may be in such an early stage of development that information about efficacy and safety is inadequate to assess risk and potential benefits. In other cases it may be unclear whether subjects have genuinely responded favorably to interventions. Under these conditions, continued access programs may not be obligatory for some early-stage studies. In some difficult cases the provision of an investigational drug that has been shown to be seriously unsafe for most patients—that is, to carry an unreasonably high level of risk—can justifiably be discontinued altogether, even if some patients have responded favorably. However, because risk and safety indexes vary significantly in subjects, what is unsafe for one group of patients may not be unduly risky for another group. A high level of risk in general therefore may not be a sufficient reason to discontinue availability to particular subjects who have responded favorably.

A Reciprocity-Based Justification of Obligations of Beneficence

Obligations of general and specific beneficence can be justified in several ways. In addition to our observations about obligations of specific beneficence based on special moral relations and roles and about the duty of rescue in particular circumstances, another justification is based on reciprocity. This approach is well suited to some areas of biomedical ethics, as we saw earlier in the discussion of expanded access. David Hume argued that the obligation to benefit others in society arises from social interactions: “All our obligations to do good to society seem to imply something reciprocal. I receive the benefits of society, and therefore ought to promote its interests.”[17] Reciprocity is the act or practice of making an appropriate, often proportional, return—for example, returning benefit with proportional benefit, countering harm-causing activities with proportional criminal sentencing, and reciprocating friendly and generous actions with gratitude. Hume’s reciprocity account rightly maintains that we incur obligations to help or benefit others, in part because we have received, will receive, or stand to receive beneficial assistance from them.

Reciprocity pervades social life. It is implausible to maintain that we are largely free of, or can free ourselves from, indebtedness to our parents, researchers in medicine and public health, and teachers. The claim that we make our way independent of our benefactors is as unrealistic as the idea that we can always act autonomously

without affecting others.[18] Codes of medical ethics have sometimes inappropriately viewed physicians as independent, self-sufficient philanthropists whose beneficence is analogous to generous acts of giving. The Hippocratic Oath states that physicians' obligations to patients represent philanthropic service, whereas obligations to their teachers represent debts incurred in the course of becoming physicians. Today many physicians and health care professionals owe a large debt to society for their formal education, training in hospitals, and the like. Many are also indebted to their patients, past and present, for learning gained from both research and practice. Because of this indebtedness, the medical profession's role of beneficent care of patients is misconstrued if modeled on philanthropy, altruism, and personal commitment. This care is rooted in a moral reciprocity of the interface of receiving and giving in return.[19]

A compelling instance of reciprocity, and one with a promising future in medicine, occurs in what the US National Academy of Medicine (NAM) calls "a learning healthcare system." A NAM report defines this type of system as "one in which knowledge generation is so embedded into the core of the practice of medicine that it is a natural outgrowth and product of the healthcare delivery process and leads to continual improvement in care."[20] A true learning health system is structured so that professionals have obligations of care to patients, and patients have specific obligations of reciprocity to facilitate learning in the health system so that care for all patients can be improved. In this institutional structure—which seems destined in the near future to increasingly become an integral part of the design of health care institutions all over the world—patients are on the receiving end of informational benefits in which the quality of their health care depends on a rapid and regular flow of information received from other patients and from other health care systems. Obligations of reciprocity call for all patients to supply information by participating in the same sort of learning activities and burdens that others have shouldered in the past to benefit them. Under these conditions, research and practice are merged in a constantly updated environment of learning designed to benefit everyone involved in the institution.

A reciprocity-based approach to beneficence has also emerged as a possible way to overcome the chronic shortage of deceased donor organs for transplantation. Appeals to obligatory or ideal beneficence to strangers have fallen far short of generating the number of organs needed to save the lives and enhance the quality of lives of patients with end-stage organ failure, many of whom die while awaiting a transplant. A reciprocity-based system would give preferential access to patients in need who previously agreed, perhaps years earlier, to donate their organs after their deaths. Declared donors' immediate family members would also be included in some proposals. In 2012, Israel became the first country to implement a reciprocity-based system.

Two models have been proposed for such programs: (1) a model of pure reciprocity restricts the pool of potential organ recipients to declared donors; (2) a model of preferential access or preferred status gives declared donors additional points toward access in an allocation point system. Both models encounter difficult questions of fairness to persons in need who were not eligible to declare their status as donors because of age or disqualifying medical conditions, but the second, nonexclusionary, preferred-status model, which Israel adopted, can handle these questions more easily. However, other justice-based moral concerns focus on how a policy might disadvantage those who are uninformed about organ donation and on how much weight should be given to the standard of declared donor status and how much to the standard of medical need.[21]

PATERNALISM: CONFLICTS BETWEEN BENEFICENCE AND RESPECT FOR AUTONOMY

The thesis that beneficence expresses a primary obligation in health care is ancient. A venerable expression appears in the Hippocratic work *Epidemics*: "As to disease, make a habit of two things—to help, or at least to do no harm."[22] Traditionally, physicians relied almost exclusively on their own judgments about their patients' needs for information and treatment. However, medicine in the modern world has increasingly encountered claims of patients' rights to receive information and to make independent judgments. As assertions of autonomy rights increased, moral problems of paternalism became clearer and more prominent.

Whether respect for the autonomy of patients should have priority over beneficence directed at those patients, that is, paternalistic beneficence, remains a central problem in clinical ethics. We will now begin to work on this problem by considering key conceptual issues.

The Nature of Paternalism

In recent biomedical ethics, paternalism has been both defended and attacked when addressing problems in clinical medicine, public health, health policy, and government policy. It is unclear in much of this literature what writers think paternalism is. The reason, we suggest, is that the notion of paternalism is a complicated and inherently contestable concept. The *Oxford English Dictionary* (*OED*) dates the term *paternalism* to the 1880s, giving its root meaning as "the principle and practice of paternal administration; government as by a father; the claim or attempt to supply the needs or to regulate the life of a nation or community in the same way a father does those of his children." This definition relies on an analogy with the father and presupposes two features of the paternal role: that the father acts beneficently (i.e., in accordance with his conception of his children's welfare interests) and that he makes all or at least some of the decisions relating to his children's welfare, rather than letting them make the decisions. In health care relationships, the analogy is that a professional has superior training, knowledge, and insight and is thus in an authoritative position to determine the patient's best interests.

Examples of paternalism in medicine include the provision of blood transfusions when patients have refused them, involuntary commitment to institutions for treatment, intervention to stop suicides, resuscitation of patients who have asked not to be resuscitated, withholding of medical information that patients have requested, denial of an innovative therapy to someone who wishes to try it, and some governmental efforts to promote health.

Paternalistic acts sometimes use forms of influence such as deception, lying, manipulation of information, nondisclosure of information, or coercion, but they may also simply involve a refusal to carry out another's wishes. According to some definitions in the literature, paternalistic actions restrict only *autonomous* choices; hence, restricting nonautonomous conduct for beneficent reasons is not paternalistic. Although one author of this book prefers this autonomy-restricted conception,[23] we here accept and refine the broader definition suggested by the *OED*: Paternalism involves an intentional nonacquiescence or intervention in another person's preferences, desires, or actions with the intention of either preventing or reducing harm to or benefiting that person. Even if a person's desires, intentional actions, and the like are not substantially autonomous, overriding them can be paternalistic under this definition.[24] For example, if a man ignorant of his fragile, life-threatening condition and sick with a raging fever attempts to leave a hospital, it is paternalistic to detain him, even if his attempt to leave does not derive from a substantially autonomous choice.

Accordingly, we define "paternalism" as "the intentional overriding of one person's preferences or actions by another person, where the person who overrides justifies the action by appeal to the goal of benefiting or of preventing or mitigating harm to the person whose preferences or actions are overridden." This definition is normatively neutral because it does not presume that paternalism is either justified or unjustified. Although the definition assumes an act of beneficence analogous to parental beneficence, it does not prejudge whether the beneficent act is justified, obligatory, misplaced, or wrong.

Problems of Medical Paternalism

Throughout the history of medical ethics, the principles of nonmaleficence and beneficence have both been invoked as a basis for paternalistic actions. For example, physicians have traditionally held that disclosing certain kinds of information can cause harm to patients under their care and that medical ethics obligates them not to cause such harm. Here is a typical case: A man brings his father, who is in his late sixties, to his physician because he suspects that his father's problems in interpreting and responding to daily events may indicate Alzheimer's disease. The man also makes an "impassioned plea" that the physician not tell his father if the tests suggest Alzheimer's. Tests subsequently indicate that the father probably does have this disease, which is a progressive brain disorder that gradually destroys memory, thinking, and abilities to carry out even simple tasks.

The physician now faces a dilemma, because of the conflict between demands of respect for autonomy, assuming that the father still has substantial autonomy and is competent at least some of the time, and demands of beneficence. The physician first considers the now recognized obligation to inform patients of a diagnosis of cancer. This obligation typically presupposes accuracy in the diagnosis, a relatively clear course of the disease, and a competent patient—none of which is clearly present in this case. The physician also notes that disclosure of Alzheimer's disease sometimes adversely affects patients' coping mechanisms, which could harm the patient, particularly by causing further decline, depression, agitation, and paranoia.[25] (See also our discussion of veracity in Chapter 8, pp. 328–34.)

Other patients—for example, those depressed or addicted to potentially harmful drugs—are unlikely to reach adequately reasoned decisions. Still other patients who are competent and deliberative may make poor choices, as judged by their physicians. When patients of either type choose harmful courses of action, some health care professionals respect autonomy by not interfering beyond attempts at persuasion, whereas others act beneficently by attempting to protect patients against the potentially harmful consequences of their own stated preferences and actions. Discussions of medical paternalism focus on how to specify or balance these principles, which principle to follow under which conditions, and how to intervene in the decisions and affairs of such patients when intervention is warranted.

Soft and Hard Paternalism

A crucial distinction exists between soft and hard paternalism.[26] In soft paternalism, an agent intervenes in the life of another person on grounds of beneficence or nonmaleficence with the goal of preventing substantially nonvoluntary conduct. Substantially nonvoluntary actions include poorly informed consent or refusal, severe depression that precludes rational deliberation, and addiction that prevents free choice and action. Hard paternalism, by contrast, involves interventions intended to prevent or mitigate harm to, or to benefit, a person, even though the person's risky choices and actions are informed, voluntary, and autonomous.

Hard paternalism usurps autonomy by either restricting the information available to a person or overriding the person's informed and voluntary choices. For example, it is an act of hard paternalism to refuse to release a competent hospital patient who will probably die outside the hospital but who requests the release in full awareness of the probable consequences. It is also an act of hard paternalism to prevent a patient capable of making reasoned judgments from receiving diagnostic information if the information would lead the patient to a state of depression. For the interventions to qualify as hard paternalism, the intended beneficiary's choices need not be fully informed or voluntary, but they must be substantially autonomous.

Soft paternalistic actions are sometimes morally complicated because of the difficulty of determining whether a person's actions are substantially nonautonomous and of determining appropriate means of protection. That we should protect persons from harm caused to them by conditions beyond their control is not controversial. Soft paternalism therefore does not involve a deep conflict between the principles of respect for autonomy and beneficence. Soft paternalism only tries to prevent the harmful consequences of a patient's actions that the patient did not choose with substantial autonomy.

This conclusion is not inconsistent with our earlier definition of paternalism as involving an intentional overriding of one person's known preferences or actions by another person. The critical matter is that some behaviors that express preferences are not autonomous. For example, some patients on medication or recovering from surgery insist that they do not want a certain physician to touch or examine them. They may be experiencing temporary hallucinations around the time of the statement. A day later they may have no idea why they stated this preference. A person's preferences can be motivated by many states and desires.

Paternalistic policies. Debates about paternalism have emerged in health policy as well as clinical ethics. Often health policies—for example, requiring a doctor's prescription for a person to acquire a type of medical device—have the goal of avoiding a harm or providing a benefit in a population in which most affected parties are not consulted about whether they agree with the policy. Policymakers understand that some percentage of the population would oppose the policy on grounds that it is autonomy depriving (by not giving them a choice),

whereas others would support the policy. In effect, the policy is intended to benefit all members of a population without consulting the autonomous preferences of all individuals, and with the knowledge that some individuals would reject the control that the policy exerts over their lives.

So-called neopaternalists or libertarian paternalists, principally coauthors Cass Sunstein and Richard Thaler, have argued for government and private institutional policies intended to protect or benefit individuals through shaping, steering, or nudging their choices in a manner that falls short of disallowing or coercing those choices.[27] In clinical care, similar arguments have supported the physician's manipulation of some patients to get them to select proper goals of care.[28] Some soft paternalists recommend policies and actions that pursue values that an intended beneficiary already, at least implicitly, holds but cannot realize because of limited capacities or limited self-control.[29] The individual's own *stated* preferences, choices, and actions are deemed unreasonable in light of *other* standards the person accepts.

By contrast, in hard paternalism the intended beneficiary does not accept the values paternalists use to determine his or her own best interests. Hard paternalism requires that the benefactor's conception of best interests prevail, and it may ban, prescribe, or regulate conduct in ways that manipulate individuals' actions to secure the benefactor's intended result. Soft paternalism, by contrast, reflects the beneficiary's conception of his or her best interests, even if the beneficiary fails to adequately understand or recognize those interests or to fully pursue them because of inadequate voluntariness, commitment, or self-control.

This conception of soft paternalism faces difficulties. Our knowledge of what an informed and competent person chooses to do is generally the *best evidence* we have of what his or her values are. For example, if a deeply religious man fails to follow the dietary restrictions of his religion, although, in the abstract, he is strongly committed to all aspects of the religion, his departures from dietary laws may be the best evidence we have of his true values on the particular matter of dietary restrictions. Because it seems correct—short of counterevidence in particular cases—that competent informed choice is the best evidence of a person's values, a justified paternalism must have adequate evidence that this assumption is misguided in a particular case.

Some prominent proponents of soft paternalism reach the conclusion that it is compatible with, rather than contrary to, autonomous choice. Sunstein and Thaler maintain that even though the idea of "libertarian paternalism" might appear to be an oxymoron, "it is both possible and desirable for private and public institutions to influence behavior while also respecting freedom of choice."[30] "Libertarian paternalism" is indeed counterintuitive, but some sense can be made of it. Suppose that available evidence were to establish that smokers psychologically discount the risks of smoking because of an "optimism bias" (among other factors). It does not follow that a government would violate their autonomy through programs intended to correct their biases—for example, through television advertisements that graphically present the suffering that often results from smoking.[31]

Libertarian paternalism builds on evidence from the cognitive sciences indicating that people have limited rationality or limited self-control that reduces their capacity to choose and act autonomously. A critical assumption is that all autonomous persons would value health over the ill health caused by smoking, and in this sense a person's deepest autonomous commitment is to be a nonsmoker. The thesis is that we are justified on autonomy grounds in arranging their choice situation in a way that likely will correct their cognitive biases and bounded rationality. However, if this position in effect holds that we should use our knowledge of cognitive biases not only to correct for failures of rationality but also to manipulate substantially autonomous people into doing what is good for them, then this position is *hard* paternalism. In short, depending on the nature of the manipulation and the nature of the affected choices, the account could turn out to be either a hard or a soft paternalism.

There is good reason for caution about libertarian paternalism.[32] The theory's supposed advantage may actually be an ethical disadvantage. This paternalism relies heavily on the thesis that there are many values that individuals would recognize or realize themselves if they did not encounter internal limits of rationality and control. The means employed, whether by health care professionals, private institutions, or governments, shape

and steer persons without thwarting their free choice. These prima facie appealing paternalistic policies and practices may face little opposition and be implemented without the transparency and publicity essential for public assessment. Paternalistic governmental policies or health care practices are susceptible to abuse if they lack transparency, public visibility, and vigorous public scrutiny.

Social norms and stigmatization. Soft paternalistic policies sometimes stigmatize conduct such as smoking. While stigmatization can change bad behavior in some contexts, it often has psychosocial costs. Proponents insist that they target *acts*, not *persons*. However, in practice, stigmatizing conduct may slide into stigmatizing people who engage in that conduct. For example, antismoking measures such as prohibitive "sin taxes" levied on cigarettes often have paternalistic goals of forcing changes in unhealthy behavior. Nevertheless, they sometimes slide from stigmatization of acts (smoking) to stigmatization of people (smokers), leading to hostility and antipathy directed at population subgroups.[33] Because smoking is now more common among lower socioeconomic groups in some countries, stigmatization thus affects socially vulnerable members of society and may involve discrimination—a matter of moral concern from the standpoint of both beneficence and justice.[34]

Soft paternalistic interventions may promote social values that eventually pave the way for hard paternalistic interventions. The history of the campaign against cigarette smoking is again instructive. It moved from disclosure of information, to sharp warnings, to soft paternalistic measures to reduce addiction-controlled unhealthy behavior, to harder paternalistic measures such as significantly increasing taxes on cigarettes.[35] In this example, paternalistic interventions remain beneficent, but they increasingly lose touch with and may even violate the principle of respect for autonomy.

The Justification of Paternalism and Antipaternalism

Three general positions appear in literature on the justification of paternalism: (1) antipaternalism, (2) paternalism that appeals to the principle of respect for autonomy as expressed through some form of consent, and (3) paternalism that appeals to the principle of beneficence. All three positions agree that some acts of soft paternalism are justified, such as preventing a man under the influence of a hallucinogenic drug from killing himself. Antipaternalists do not object to such interventions because substantially autonomous actions are not at stake.

Antipaternalism. Antipaternalists oppose hard paternalistic interventions for several reasons. One motivating concern focuses on the potential adverse consequences of giving paternalistic authority to the state or to a group such as physicians. Antipaternalists regard rightful authority as residing in the individual. The argument for this position rests on the principle of respect for autonomy as discussed in Chapter 4 (pp. 99–106): Hard paternalistic interventions display disrespect toward autonomous agents and fail to treat them as moral equals, treating them instead as less-than-independent determiners of their own good. If others impose their conception of the good on us, they deny us the respect they owe us, even if they have a better conception of our needs than we do.[36]

Antipaternalists also hold that paternalistic standards are too broad and authorize and institutionalize too much intervention when made the basis of policies. If this charge is correct, paternalism allows an unacceptable latitude of judgment. Consider the example of a sixty-five-year-old man who has donated a kidney to one of his sons and now volunteers to donate his second kidney when another son needs a transplant, an act most would think not in his best interests even though he contends that he could survive on dialysis. Are we to commend him, ignore him, or deny his request? Hard paternalism suggests that it would be permissible and perhaps obligatory to stop him or at least to refuse to carry out his request, a judgment that could easily be made a matter of institutional or public policy. If so, antipaternalists argue, the state or an institution is permitted, in principle, to prevent its morally heroic citizens from acting in a manner "harmful" to themselves.

However, some interventions that are paternalistic (in our broad understanding of paternalism) can be accepted by antipaternalists. A medical example with an extensive antipaternalistic literature is the involuntary hospitalization of persons who have neither been harmed by others nor actually harmed themselves, but who have been assessed as at risk of such harm because of a documented disorder that substantially compromises

their autonomous choices. In this case, a double paternalism is common—a paternalistic justification for both commitment and forced therapy. Antipaternalists could regard this kind of intervention as justified by the intent to benefit, emphasizing that beneficence does not here conflict with respect for autonomy because the intended beneficiary lacks substantial autonomy.

Paternalism justified by consent. Some appeal to *consent* to justify paternalistic interventions—be it rational consent, subsequent consent, hypothetical consent, or some other type of consent. As Gerald Dworkin states it, "The basic notion of consent is important and seems to me the only acceptable way to try to delimit an area of justified paternalism." Paternalism, he maintains, is a "social insurance policy" to which fully rational persons would subscribe in order to protect themselves.[37] They would know, for example, that they might be tempted at times to make decisions that are far-reaching, potentially dangerous, and irreversible. At other times, they might suffer irresistible psychological or social pressures to take actions that are unreasonably risky. In still other cases, persons might not sufficiently understand the dangers of their actions, such as medical facts about the effects of smoking, although they might believe that they have a sufficient understanding. Those who use consent as a justification conclude that, as fully rational persons, we would consent to a limited authorization for others to control our actions if our autonomy becomes defective or we are unable to make the prudent decision that we otherwise would make.[38]

A theory that appeals to rational consent to justify paternalistic interventions has attractive features, particularly its attempt to harmonize principles of beneficence and respect for autonomy. However, this approach does not incorporate an individual's actual consent and is therefore not truly consent based. It is best to keep autonomy-based justifications at arm's length from both paternalism and hypothetical, rational-persons arguments. Beneficence alone justifies truly paternalistic actions, exactly as it justifies parental actions that override children's preferences.[39] Children are controlled not because we believe they will subsequently consent to or rationally approve our interventions. We control them because we believe they will have better, or at least less dangerous, lives.

Paternalism justified by prospective benefit. Accordingly, the justification of paternalistic actions that we recommend places benefit on a scale with autonomy interests and balances both: As a person's interests in autonomy increase and the benefits for that person decrease, the justification of paternalistic action becomes less plausible; conversely, as the benefits for a person increase and that person's autonomy interests decrease, the justification of paternalistic action becomes more plausible. Preventing minor harms or providing minor benefits while deeply disrespecting autonomy lacks plausible justification, but actions that prevent major harms or provide major benefits while negatively affecting (or "disrespecting") autonomy in only minor ways have a plausible paternalistic rationale. As we will now argue, even hard paternalistic actions can under some conditions be justified on these grounds.[40]

Justified hard paternalism. An illustrative (and actual) case provides a good starting point for reflection on the conditions of justified hard paternalism: A physician obtains the results of a myelogram (a graph of the spinal region) following examination of a patient. The test yields inconclusive results and needs to be repeated, but it also suggests a serious pathology. When the patient asks about the test results, the physician decides on grounds of beneficence to withhold potentially negative information, knowing that, on disclosure, the patient will be distressed and anxious. Based on her experience with other patients and her ten-year knowledge of this particular patient, the physician is confident that the information would not affect the patient's decision to consent to another myelogram. Her sole motivation in withholding the information is to spare the patient the emotional distress of processing negative but not fully confirmed information, which, at this time, seems premature and unnecessary. However, the physician intends to be completely truthful with the patient about the results of the second test and intends to disclose the information well before the patient would need to decide about surgery. This physician's act of temporary nondisclosure is morally justified because she has determined that beneficence has temporary priority over respect for autonomy.[41] Such minor hard paternalistic actions are common in medical practice and in our view are sometimes warranted.

To consolidate the discussion thus far, hard paternalism involving a health professional's intervention is justified only if the following conditions are satisfied (see further our conditions for constrained balancing in Chapter 1, pp. 22–24):

1. 1. A patient is at risk of a significant, preventable harm or failure to receive a benefit.
2. 2. The paternalistic action will probably prevent the harm or secure the benefit.
3. 3. The intervention to prevent harm to or to secure a benefit for the patient probably outweighs the risks to the patient of the action taken.
4. 4. There is no morally better alternative to the limitation of autonomy that will occur.
5. 5. The least autonomy-restrictive alternative that will prevent the harm or secure the benefit is adopted.

A sixth condition could be added requiring that a paternalistic action not damage *substantial* autonomy interests, as would occur if one were to override the decision of a Jehovah's Witness patient who, from deep conviction, refuses a blood transfusion. To intervene forcefully by providing the transfusion would substantially infringe the patient's autonomy and could not be justified under this additional condition. However, some cases of justified hard paternalism do cross the line of minimal infringement. In general, as the risk to a patient's welfare increases or the likelihood of an irreversible harm increases, the likelihood of a justified paternalistic intervention correspondingly increases.

The following case plausibly supports a hard paternalistic intervention despite the fact that it involves more than minimal infringement of respect for autonomy: A psychiatrist is treating a patient who is sane but who acts in what appear to be bizarre ways. He is acting conscientiously on his unique religious views. He asks a psychiatrist a question about his condition, a question that has a definite answer but which, if answered, would lead the patient to engage in seriously self-maiming behavior such as plucking out his right eye to fulfill what he believes to be his religion's demands. Here the doctor acts paternalistically, and justifiably, by concealing information from this patient, who is rational and otherwise informed. Because the infringement of the principle of respect for autonomy is more than minimal in this case (the stated religious views being central to this patient's life plan), a sixth condition requiring no substantial infringement of autonomy is not a necessary condition of all cases of justified hard paternalism.

Problems of Suicide Intervention

The tenth leading cause of death in the United States as this book goes to press is suicide. In 2016, nearly 45,000 persons committed suicide, an increase of roughly 30% since 1999. Data available from about half of the US states indicate that over 50% of persons committing suicide were not known to have mental health problems.[42] These striking figures suggest that improvements in beneficence-based suicide prevention programs have not been as effective as planners of these programs anticipated.

We will focus on suicide intervention, that is, interventions with the intent to prevent suicides. The state, religious institutions, and health care professionals have traditionally asserted jurisdiction to intervene in suicide attempts. Those who intervene do not always justify their actions on paternalistic grounds, but paternalism has been a common justification.

However, several conceptual questions about the term *suicide* make it difficult to categorize some acts as suicides.[43] A classic example of these difficulties involves Barney Clark, who became the first human to receive an artificial heart. He was given a key to use to turn off the compressor if he decided he wanted to die. As Dr. Willem Kolff noted, perhaps from an antipaternalistic perspective, if the patient "suffers and feels it isn't worth it any more, he has a key that he can apply. … I think it is entirely legitimate that this man whose life has been extended should have the right to cut it off if he doesn't want it, if [his] life ceases to be enjoyable."[44]

Would Clark's use of the key to turn off the artificial heart have been an act of suicide? If he had refused to accept the artificial heart in the first place, few would have labeled his act a suicide. His overall condition was extremely poor, the artificial heart was experimental, and no suicidal intention was evident. If, on the other hand,

Clark had intentionally shot himself with a pistol while on the artificial heart, his act would have been classified as suicide.

Our main concern is paternalistic intervention in acts of attempted suicide. The primary moral issue is that if autonomous suicide is a protected moral right, then the state, health professionals, and others have no legitimate grounds for intervention in autonomous suicide attempts. No one doubts that we should intervene to prevent suicide by substantially nonautonomous persons, and few wish to return to the days when suicide was a criminal act. However, if there is an autonomy right to commit suicide, then we could not legitimately attempt to prevent an autonomous but imprudent individual from committing suicide.

A clear and relevant example of attempted suicide appears in the following case, involving John K., a thirty-two-year-old lawyer. Two neurologists independently confirmed that his facial twitching, which had been evident for three months, was an early sign of Huntington's disease, a neurological disorder that progressively worsens, leads to irreversible dementia, and is uniformly fatal in approximately ten years. His mother suffered a horrible death from the same disease, and John K. had often said that he would prefer to die than to suffer the way his mother had suffered. Over several years he was anxious, drank heavily, and sought psychiatric help for intermittent depression. After he received this diagnosis, he told his psychiatrist about his situation and asked for help in committing suicide. After the psychiatrist refused to help, John K. attempted to take his own life by ingesting his antidepressant medication, leaving a note of explanation to his wife and child.[45]

Several interventions occurred or could have occurred in this case. First, the psychiatrist refused to assist John K.'s suicide and would have sought involuntary commitment had John K. not insisted, convincingly, that he did not plan to kill himself anytime soon. The psychiatrist appears to have thought that he could provide appropriate psychotherapy over time. Second, John K.'s wife found him unconscious and rushed him to the emergency room. Third, the emergency room staff decided to treat him despite the suicide note. The question is which, if any, of these possible or actual interventions is justifiable.

A widely accepted account of our obligations relies on a strategy of *temporary* intervention devised by John Stuart Mill. On this account, provisional intervention is justified to ascertain whether a person is acting autonomously, but further intervention is unjustified once it is clear that the person's actions are substantially autonomous. Glanville Williams used this strategy in a classic statement of the position:

> If one suddenly comes upon another person attempting suicide, the natural and humane thing to do is to try to stop him, for the purpose of ascertaining the cause of his distress and attempting to remedy it, or else of attempting moral dissuasion if it seems that the act of suicide shows lack of consideration for others, or else again from the purpose of trying to persuade him to accept psychiatric help if this seems to be called for. … But nothing longer than a temporary restraint could be defended. I would gravely doubt whether a suicide attempt should be a factor leading to a diagnosis of psychosis or to compulsory admission to a hospital. Psychiatrists are too ready to assume that an attempt to commit suicide is the act of mentally sick persons.[46]

This strong antipaternalist stance might be challenged on two grounds. First, failure to intervene in a more forceful manner than Williams allows symbolically communicates to potentially suicidal persons a lack of communal concern and seems to diminish communal responsibility. Second, many persons who commit or attempt to commit suicide are mentally ill, clinically depressed, or destabilized by a crisis and are, therefore, not acting autonomously. Many mental health professionals believe that suicides generally result from maladaptive attitudes or illnesses needing therapeutic attention and social support. In a typical circumstance the suicidal person plans how to end life while simultaneously holding fantasies about how rescue will occur, including rescue from the negative circumstances prompting the suicide as well as rescue from the suicide itself. If the suicide springs from clinical depression for which the patient has sought treatment or constitutes a call for help, a failure to intervene shows disrespect for the person's deepest autonomous wishes, including his or her hopes for the future.

Nonetheless, caution is needed in any such account of communal beneficence, which may be expressed paternalistically through unjustifiably forceful interventions. Although suicide has been decriminalized in most countries, a suicide attempt, irrespective of motive, almost universally provides a legal basis for public officers to intervene, as well as grounds for at least temporary involuntary hospitalization.[47] Still, the burden of proof falls on those who claim that the patient's judgment is insufficiently autonomous.

Consider the following instructive example involving Ida Rollin, seventy-four years old and suffering from ovarian cancer. Her physicians truthfully told her that she had only a few months to live and that her dying would be painful and upsetting. Rollin indicated to her daughter that she wanted to end her life and requested assistance. The daughter secured some pills and conveyed a doctor's instructions about how they should be taken. When the daughter later expressed reservations about these plans, her husband reminded her that they "weren't driving, she [Ida Rollin] was," and that they were only "navigators."[48]

This metaphor-laden reference to rightful authority is a reminder that those who propose suicide intervention to prevent such persons from control over their lives need a moral justification that fits the context. Occasions arise in health care and beyond when it is appropriate to step aside and allow a person to bring his or her life to an end, and perhaps even to assist in facilitating the death, just as occasions exist when it is appropriate to intervene. (See Chapter 5 on physician-assisted forms of ending life, pp. 184–93.)

Denying Requests for Nonbeneficial Procedures

Patients and surrogates sometimes request medical procedures that the clinician is convinced will not be beneficial and will perhaps be harmful. Sometimes denials of such requests are paternalistic.

Passive paternalism. A passive paternalistic act occurs when professionals refuse, for reasons of beneficence, to carry out a patient's positive preferences for an intervention.[49] The following is a case in point: Elizabeth Stanley, a sexually active twenty-six-year-old intern, requests a tubal ligation, insisting that she has thought about this request for months, dislikes available contraceptives, does not want children, and understands that tubal ligation is irreversible. When the gynecologist suggests that she might someday want to get married and have children, she responds that she would either find a husband who did not want children or adopt children. She thinks that she will not change her mind and wants the tubal ligation to make it impossible for her to reconsider. She has scheduled vacation time from work in two weeks and wants the surgery then.[50]

If a physician refuses to perform the tubal ligation on grounds of the patient's benefit, the decision is paternalistic. However, if the physician refuses purely on grounds of conscience ("I won't do such procedures as a matter of personal moral convictions"), the refusal may not rest on any type of paternalistic motive.

Passive paternalism is usually easier to justify than active paternalism, because physicians generally do not have a moral obligation to carry out their patients' wishes when they are incompatible with accepted standards of medical practice, conflict with their physician's judgment about medical benefit or harm, or are against the physician's conscience. Each type of passive paternalism may be justified in some cases, but not in others.

Medical futility. Passive paternalism appears in decisions not to provide patient-requested procedures that are deemed medically futile. (We treated the topic of medical futility in Chapter 5, pp. 171–74). Consider the classic case of eighty-five-year-old Helga Wanglie, who was maintained on a respirator in a persistent vegetative state. The hospital sought to stop the respirator on grounds that it was nonbeneficial in that it could not heal her lungs, palliate her suffering, or enable her to experience the benefits of life. Surrogate decision makers—her husband, a son, and a daughter—wanted life support continued on grounds that Mrs. Wanglie would not be better off dead, that a miracle could occur, that physicians should not play God, and that efforts to remove her life support epitomize "moral decay in our civilization."[51] (Because the request for continued treatment came from the family rather than the patient, it can be viewed as a case of passive paternalism only on the assumption that the family is asserting what it takes to be Mrs. Wanglie's wishes.)

If life support for such patients truly is futile, denying patients' or surrogates' requests for treatment is warranted. In these circumstances "clinically nonbeneficial interventions" may be preferable to the term *futility*.[52] Typically a claim of futility is not that an intervention will harm the patient in violation of the principle of nonmaleficence but that it will not produce the benefit the patient or the surrogate seeks. A justified belief in futility cancels a professional's obligation to provide a medical procedure. However, it is not clear that the language of futility illuminates the range of relevant ethical issues in passive paternalism, in part because of its vague uses, which we discussed in Chapter 5 (where we argued that, for all its problems, "futile" is still superior to a recently proposed substitution of the even vaguer term "inappropriate." See pp. 172–73.)[53]

BALANCING BENEFITS, COSTS, AND RISKS

Thus far we have concentrated on the role of the principle of beneficence in clinical medicine, health care, and public policy. We will now consider how principles of beneficence, particularly the principle of utility in our sense of the term (see pp. 217–18) can be applied to health policies through tools that analyze and assess benefits relative to costs and risks. Because formal analysis has assumed a critical role in policy decision making, the importance of ethical assessment of these methods has increased. These tools often are morally unobjectionable and may even be morally required in some circumstances, but problems do attend their use.

Physicians routinely base judgments about the most suitable medical treatments on the balance of probable benefits and probable harms for patients. This criterion is also used in judgments about the ethical acceptability of research involving human subjects. These judgments consider whether the probable overall benefits—for society as well as subjects—outweigh the risks to subjects. In submitting a research protocol involving human subjects to an institutional review board (IRB) for approval, an investigator is expected to array the risks to subjects and probable benefits to both subjects and society, and then to explain why the probable benefits outweigh the risks. When IRBs array risks and benefits, determine their respective weights, and reach decisions, they typically use informal techniques such as expert judgments based on reliable data and analogical reasoning based on precedents. We focus here on techniques that employ formal, quantitative analysis of costs, risks, and benefits and offer an ethical assessment of their use as ways of applying principles of beneficence.

The Nature of Costs, Risks, and Benefits

We start with some basic conceptual questions about costs, risks, and benefits. *Costs* include the resources required to bring about a benefit as well as the negative effects of pursuing and realizing that benefit. We concentrate on costs expressed in monetary terms—the primary interpretation of costs in cost-benefit and cost-effectiveness analysis. The term *risk*, by contrast, refers to a possible future harm, where harm is defined as a setback to interests, particularly in life, health, or welfare. Expressions such as *minimal risk*, *reasonable risk*, and *high risk* usually refer to the chance of a harm's occurrence—its probability—but often also to the severity of the harm if it occurs—its magnitude.[54]

Statements of risk are *descriptive* inasmuch as they state the probability that harmful events will occur. They are *evaluative* inasmuch as they attach a value to the occurrence or prevention of these events. Statements of risk presume a prior negative evaluation of some condition. At its core, a circumstance of risk involves a possible occurrence of something that has been evaluated as harmful along with an uncertainty about its actual occurrence that can be expressed in terms of its probability. Several types of risk exist, including physical, psychological, financial, and legal risks.

The term *benefit* sometimes refers to cost avoidance and risk reduction, but more commonly in biomedicine it refers to something of positive value, such as life or improvement in health. Unlike *risk*, *benefit* is not itself a probabilistic term. *Probable benefit* is the proper contrast to risk, and benefits are comparable to harms rather than to risks of harm. Accordingly, risk-benefit relations are best understood in terms of a ratio between the probability and magnitude of an anticipated benefit and the probability and magnitude of an anticipated harm.

Risk Assessment and Values in Conflict

Risk *assessment* involves the analysis and evaluation of probabilities of negative outcomes, especially harms. Risk *identification* seeks to locate a hazard. Risk *estimation* determines the probability and magnitude of harm from that hazard. Risk *evaluation* determines the acceptability of the identified and estimated risks, often in relation to other objectives. Evaluation of risk in relation to probable benefits is often labeled *risk-benefit analysis* (RBA), which may be formulated in terms of a ratio of expected benefits to risks and may lead to a judgment about the acceptability of the risk under assessment. Risk identification, estimation, and evaluation are all stages in risk assessment. The next stage in the process is risk *management*—the set of individual, institutional, or policy responses to the analysis and assessment of risk, including decisions to reduce or control risks.[55] For example, risk management in hospitals includes setting policies to reduce the risk of medical malpractice suits as well as the risk of accidents, injuries, and medical errors.

Risk assessment informs technology assessment, environmental impact statements, and public policies protecting health and safety. The following schema of magnitude and probability of harm captures important features of risk assessment:

		Magnitude of Harm	
		Major	*Minor*
Probability of Harm	*High*	1	2
	Low	3	4

Under category 4, questions arise about whether some risks are so insignificant, in terms of either probability or magnitude of harm or both, as not to merit attention. So-called *de minimis* risks are acceptable risks because they can be interpreted as effectively zero. According to the FDA, a risk of less than one cancer per million persons exposed is *de minimis*. However, using this quantitative threshold or cutoff point in a *de minimis* approach may be problematic. For instance, an annual risk of one cancer per million persons for the US population would produce the same number of fatalities (i.e., 300) as a risk of one per one hundred in a town with a population of 30,000. In focusing on the annual risk of cancer or death to one individual per million, the *de minimis* approach may neglect the cumulative, overall level of risk created for individuals over their lifetimes by the addition of several one-per-million risks.[56]

Risk assessment also focuses on the acceptability of risks relative to the benefits sought. With the possible exception of *de minimis* risks, most risks will be considered acceptable or unacceptable in relation to the probable benefits of the actions that carry those risks—for example, the benefits of radiation, hormone therapy, or a surgical procedure in managing prostate cancer or the benefits of nuclear power or toxic chemicals in the workplace.[57] Vigorous disputes sometimes emerge over competing risk-benefit analyses. Consider, for instance, the judgments about newborn male circumcision by two well-informed medical societies: The Canadian Pediatrics Society concluded that the benefits of circumcision do not in most situations exceed its risks, whereas the American Society of Pediatrics (as well as the US Centers for Disease Control) held that its benefits exceed its risks, resulting in divergent recommendations to parents.[58]

Risk-benefit analyses in the regulation of drugs and medical devices. Some of the conceptual, normative, and empirical issues in risk assessment and in RBA are evident in governmental regulation of drugs and medical devices.

The US FDA requires three phases of human trials of drugs prior to regulatory approval. Each stage involves RBA to determine whether to proceed to the next stage and whether to approve a drug for wider use. As noted above, patients, physicians, and other health care professionals have often criticized the process of drug approval because of the length of time required. Some critics contend that the standard of evidence for a favorable risk-

benefit ratio is too high and thus severely limits patients' access to promising new drugs, often in times of dire need created by serious, even fatal, medical conditions. (See the discussion of expanded access earlier in this chapter, pp 225–27.) Other critics charge that the process is not rigorous enough in view of the problems that sometimes appear after drug approval.[59] A related and morally important criticism is that approved drugs that turn out to be inefficacious or unsafe in wider use are sometimes not removed from the market quickly enough, if at all. FDA policy is that drugs are removed from the market whenever their risks outweigh their benefits. For example, a drug might be removed from the market because of an uncorrectable safety issue that was unknown at the point of approval. However, removal from the market may not occur for many years after it becomes reasonably clear that risks outweigh benefits.

An example involving medical devices presents a classic case of difficult and controversial RBAs and assessments undertaken by the FDA in its regulatory decisions. For more than thirty years, thousands of women used silicone gel-filled breast implants to augment their breast sizes, to reshape their breasts, or to reconstruct their breasts following mastectomies for cancer or other surgery. (Saline-filled implants were also used. Both types have a silicone outer shell, but the silicone gel-filled implants generated greater concern.) These implants were already on the market when legislation in 1976 required that manufacturers provide data about safety and efficacy for certain medical devices. Implant manufacturers were not required to provide these data unless questions arose. The health and safety concerns that subsequently emerged centered on the silicone gel-filled implants' longevity, rate of rupture, and link with various diseases.

Defenders of the complete prohibition of these implants contended that no woman should be allowed to take a risk of unknown but potentially serious magnitude because her consent might not be adequately informed. FDA Commissioner David Kessler and others defended a restrictive policy, which was implemented in 1992. Kessler argued that for "patients with cancer and others with a need for breast reconstruction," a favorable risk-benefit ratio could exist in carefully controlled circumstances.[60] Sharply distinguishing candidates for reconstruction following surgery from candidates for augmentation, he held that a favorable risk-benefit ratio existed only for candidates for reconstruction.

Because candidates for augmentation still had breast tissue, they were considered to be at "higher risk" from these implants. In the presence of an implant, the argument went, mammography might not detect breast cancer, and the use of mammography could create a risk of radiation exposure in healthy young women with breast tissue who have silent ruptures of the silicone gel-filled implant without symptoms. Kessler wrote: "In our opinion the risk-benefit ratio does not at this time favor the unrestricted use of silicone gel breast implants in healthy women."

Kessler denied that this decision involved "any judgment about values," but critics rightly charged that, in fact, it was based on contested values and was inappropriately paternalistic. There is evidence that the FDA gave an unduly heavy weight to unknown risks largely because the agency discounted the self-perceived benefits of breast implants for women except in cases of reconstruction. The agency then held these implants to a high standard of safety instead of allowing women to decide for themselves whether to accept the risks for their own subjective benefits.[61]

If the evidence had indicated high risk relative to benefit, as well as unreasonable risk-taking by women, a different conclusion might have been warranted, but evidence available at the time and since points in the other direction. The FDA policy was unjustifiably paternalistic, noticeably so when compared to the less restrictive public decisions reached in European countries.[62] A more defensible, nonpaternalistic policy would have permitted the continued use of silicone gel-filled breast implants, regardless of the users' biological conditions and aims, while requiring adequate disclosure of information about risks. Raising the level of disclosure standards, as the FDA has done in some cases, would have been more appropriate than restraining choice.

In 2006, based on new data from manufacturers and assessments by its advisory committees, the FDA approved the marketing of two companies' silicone gel-filled breast implants to women of all ages for breast reconstruction and to women twenty-two years old and older for breast augmentation.[63] Even though these breast implants have "frequent local complications and adverse outcomes," the FDA determined that their

benefits and risks are "sufficiently well understood for women to make informed decisions about their use,"[64] a conclusion that allows the FDA to escape the problems of paternalism that plagued earlier policies. The FDA has since continued to monitor data about implants and communicate new safety information. It has also called for manufacturers and physicians to provide current and balanced information to help inform women's decisions.

Another concern is that the RBA used in the approval and regulation of drugs or devices is sometimes too narrow or limited. For example, as this edition of our book goes to press, the United States faces a devastating opioid epidemic that is much worse than in most other countries. At least two million people in the United States have an opioid use disorder (OUD), including dependence and abuse, that involves prescribed medications. Six hundred thousand more have an OUD that involves heroin. Approximately ninety people die each day from an opioid overdose. This epidemic has resulted in part from important, badly needed, and overdue beneficent efforts to treat patients' pain more effectively. In light of the epidemic's wide-ranging individual and societal harms and costs, a consensus committee of the US National Academies called on the FDA to use a broader analysis of risks and benefits in approving and monitoring prescription opioids for pain management.[65] This analysis is broader than usual in at least two ways: It involves a comprehensive, systematic public health evaluation and a more thorough post-approval monitoring and oversight that attends to patterns of prescription and use.

The FDA's approach to drug approval usually focuses specifically on the product, the drug, in light of the data that the manufacturer generates and provides on that drug. Then the FDA balances the probable benefits indicated by these data against the risks that are known, or unknown, at the time of the analysis. However, this approach may fail to balance adequately the individual and societal benefits and risks of opioid drugs as they are actually prescribed and used in practice, where they produce a variety of effects on households and society generally. It is important, but insufficient, to evaluate the probable benefits (including relief of pain and improvement of function) and risks (including respiratory depression and death as well as opioid use disorder) for individual patients. It is also necessary to evaluate the benefits and risks to others in a patient's household and in the community, such as effects on crime and unemployment, along with the drug's potential impact on legal and illegal markets for opioids, diversion of prescription opioids, transition to illicit opioids, and injection-related harms such as HIV and hepatitis C virus. Moreover, the consensus committee's report called for attention to distinctive benefit-risk profiles of different subpopulations and geographic areas—a concern of equity.[66] In short, broad public health considerations need to be incorporated thoroughly and systematically into regulatory decisions about opioid approval.

Because this task is so broad, incorporates so many factors and variables, and requires data that are difficult to obtain in high quality, a formal, comprehensive, systematic RBA will be difficult and perhaps impossible to achieve. More likely the FDA along with other involved public and private bodies will need to balance, in formal and informal ways, the benefits and risks to both patients needing pain relief and others exposed to the wide-ranging risks in order to determine appropriate policies. This balancing should occur in a transparent, public, deliberative context, with input from all affected stakeholders.

We reach two general conclusions: First, it is morally legitimate and often obligatory for society to act beneficently through the government and its agencies to protect citizens from medical drugs and devices that are harmful or that have not been established to be safe and efficacious. Hence, the FDA and comparable agencies play a justifiable regulatory role. Our conclusion that the FDA should not have severely restricted or prohibited the use of silicone gel-filled breast implants should not be interpreted as an argument against the agency's indispensable social role. As the opioid epidemic indicates, the conduct and use of RBA in drug and device approval and monitoring may need to be broader than often thought even though it may inevitably be less formal and less systematic than desired because of the wide range of potentially relevant factors. Second, RBAs are not value-free. Values are evident in various RBA-based decisions, including those made in the breast implant case and in the evaluation of opioid drugs.

Risk perception. Perceptions of risk vary in different human communities, and an individual's perception of risks in any of these communities may differ from an expert's assessment. Variations may reflect different goals and "risk budgets," as well as different qualitative assessments of particular risks, including whether the risks in question are voluntary, controllable, highly salient, novel, or dreaded.[67]

Differences in risk perception suggest some limits of attempts to use quantitative statements of probability and magnitude in reaching conclusions about the acceptability of different risks. The public's informed but subjectively interpreted perception of a possible or probable harm should be considered and given substantial weight when public policy is formulated, but the appropriate weighting will vary with each case. The public sometimes holds factually mistaken or only partially informed views about risks that experts can identify. These mistaken or underinformed public views can and should be corrected through a fair public policy process.[68]

Precaution: principle or process? Sometimes a new technology such as nanotechnology or a novel activity such as injecting bovine growth hormone into dairy cows appears to pose a health threat or create a hazard, thereby evoking public concern. Scientists may lack evidence to determine the magnitude of the possible negative outcome or the probabilities of its occurrence, perhaps because of uncertain cause-effect relations. The risks cannot be quantified and an appropriate benefit-risk-cost analysis is not constructible. At best, beneficence can be implemented through *precautionary* measures. Which actions, if any, are justifiable in the face of uncertain risks?

Several common maxims come to mind: Better safe than sorry; look before you leap; and an ounce of prevention is worth a pound of cure. As rough guides for decision making, these maxims are unobjectionable. A so-called precautionary principle has been implemented in some international treaties as well as in laws and regulations in several countries to protect the environment and public health.[69] It is misleading to speak, as some commentators and policies do, about *the* precautionary principle because there are so many different versions of the concept of precaution in law and policy and of proposed normative principles that have different strengths and weaknesses. One analysis identifies as many as nineteen different formulations,[70] and views expressed about particular precautionary measures are rarely expressed in a form that is truly a *principle*.

A precautionary principle, in its most demanding versions, could be a recipe for paralysis; it may be too abstract to give substantive, practical guidance, and appeals to it may lead parties to carefully examine only one narrow set of risks while ignoring other risks and potential benefits.[71] For example, appealing to this principle to prevent scientific research using human cells and animal chimeras, because of a perceived but vague risk of adverse consequences, may neglect significant potential health benefits that could result from the research. Precaution often has a price.[72] Perils created by some formulations and uses of a precautionary principle include distortion of public policy as a result of speculative and theoretical threats that divert attention from real, albeit less dramatic, threats.

However, if properly formulated, some precautionary approaches, processes, and measures are meaningful and justified.[73] Depending on what is valued and what is at risk, it may be ethically justifiable and even obligatory to take steps, in the absence of conclusive scientific evidence, to avoid a hazard where the harm would be both serious and irreversible—that is, a catastrophe.[74] Triggering conditions for these measures include plausible evidence of potential major harm where it is not possible to adequately characterize and quantify risk because of scientific uncertainty and ignorance. The process of developing precautionary norms should not be viewed as an alternative to risk analysis and scientific research. It should instead be viewed as a way to supplement risk appraisals when the available scientific evidence does not permit firm characterizations of the probability or magnitude of plausible risks.

Prudent use of precaution is more an approach or a process than an action based on a genuine principle, and it needs to be justified by a rigorous interpretation of the principles of beneficence and nonmaleficence. "We do not need a precautionary principle," Christian Munthe writes; "we need a policy that expresses a proper degree of precaution."[75] Measures commonly associated with a precautionary process include transparency, involvement of the public, and consultation with experts about possible responses to threats marked by ignorance or uncertainty about probabilities and magnitudes. Although transparency sometimes heightens fears, the public good is best served by risk-avoidance or risk-reduction policies that are generally consistent with the society's basic values and the public's reflective preferences. The acceptance or rejection of any particular precautionary approach will depend on a careful weighing of ethical, social, cultural, and psychological considerations.[76]

It is easy to oversimplify and unduly magnify cultural differences by suggesting, for instance, that Europe is more precaution-oriented than the United States. Even if precautionary approaches may have more traction in laws, regulations, and discourse in Europe than in the United States, both adopt a variety of precautionary measures in response to the same and to different perceived threats or hazards.[77]

Cost-Effectiveness and Cost-Benefit Analyses

Cost-effectiveness analysis (CEA) and cost-benefit analysis (CBA) are widely used, but sometimes controversial, tools of formal analysis underlying public policies regarding health, safety, and medical technologies.[78] Some policies are directed at burgeoning demands for expensive medical care and the need to contain costs. In assessing such policies, CEA and CBA appear precise and helpful because they present trade-offs in quantified terms.[79] Yet they are not unproblematic.

Defenders of these techniques praise them as ways to reduce the intuitive weighing of options and to avoid subjective and political decisions. Critics claim that these methods of analysis are not sufficiently comprehensive, that they fail to include all relevant values and options, that they frequently conflict with principles of justice, and that they are often themselves subjective and biased. Critics also charge that these techniques concentrate decision-making authority in the hands of narrow, technical professionals (e.g., some health economists) who often fail to understand moral, social, legal, and political constraints that legitimately limit use of these methods.

CEA and CBA use different terms to state the value of outcomes. CBA measures both the benefits and the costs in monetary terms, whereas CEA measures the benefits in nonmonetary terms, such as years of life, quality-adjusted life-years, or cases of disease. CEA offers a bottom line such as “cost per year of life saved,” whereas CBA offers a bottom line of a benefit-cost ratio stated in monetary figures that express the common measurement. Although CBA often begins by measuring different quantitative units (such as number of accidents, statistical deaths, and number of persons treated), it attempts to convert and express these seemingly incommensurable units of measurement into a common figure.

Because it uses the common metric of money, CBA in theory permits a comparison of programs that save lives with, for example, programs that reduce disability or accomplish other goals, such as public education. By contrast, CEA does not permit an evaluation of the inherent worth of programs or a comparative evaluation of programs with different aims. Instead, CEA functions best to compare and evaluate different programs that share an identical aim, such as saving years of life.

Many CEAs involve comparing alternative courses of action that have similar health benefits to determine which is the most cost-effective. A simple and now classic example is the use of the guaiac test, an inexpensive test for detecting minute amounts of blood in the stool. Such blood may result from several problems, including hemorrhoids, benign intestinal polyps, or colonic cancer. A guaiac test cannot identify the cause of the bleeding, but if there is a positive stool guaiac and no other obvious cause for the bleeding, physicians undertake other tests. In the mid-1970s, the American Cancer Society proposed using six sequential stool guaiac tests to screen for colorectal cancers. Two analysts prepared a careful CEA of the six stool guaiac tests. They assumed that the initial test costs four dollars, that each additional test costs one dollar, and that each successive test detects many fewer cases of cancer. They then determined that the marginal cost *per case of detected cancer* increased dramatically: $1,175 for one test; $5,492 for two tests; $49,150 for three tests; $469,534 for four tests; $4.7 million for five tests; and $47 million for the full six-test screen.[80] Such findings do not dictate a conclusion, but the analysis provides relevant data for a society needing to allocate resources, for insurance companies and hospitals setting policies, for physicians making recommendations to patients, and for patients considering diagnostic procedures.

However, confusion can mar the conduct and uses of CEA. In some cases, when two programs are compared, the cost savings of one may be sufficient to view it as more cost-effective than the other; but we should not confuse CEA with either reduced costs or increased effectiveness, because the best conclusions often depend on

both together. A program may be more cost-effective than another even if (1) it costs more, because it may increase medical effectiveness, or (2) it leads to an overall decrease in medical effectiveness, because it may greatly reduce the costs. No form of analysis has the moral power to dictate the use of a particular medical procedure simply because that procedure has the lowest cost-effectiveness ratio. To assign priority to the alternative with the lowest cost-effectiveness ratio is to view medical diagnosis and therapy in unjustifiably narrow terms.

THE VALUE AND QUALITY OF LIFE

We turn now to controversies regarding how to place a value on a human life, which have centered on CBAs, and to controversies over the value of quality-adjusted life-years (QALYs), which have centered on CEAs.

Valuing Lives

We begin by considering indicators of appropriate social beneficence that involve assigning an economic value to human life. A society may spend amount *x* to save a life in one setting (e.g., by reducing the risk of death from cancer) but only spend amount *y* to save a life in another setting (e.g., by reducing the risk of death from mining accidents). One objective in determining the value of a life is to develop consistency across practices and policies.

Analysts have developed several methods to determine the value of human life. These include discounted future earnings (DFE) and willingness to pay (WTP). According to DFE, we can determine the monetary value of lives by considering what people at risk of some disease or accident could be expected to earn if they survived. Although this approach can help measure the costs of diseases, accidents, and death, it risks reducing people's value to their potential economic worth and gives an unfair priority to those who would be expected to have greater future earnings.

WTP, which is now more commonly used, considers how much individuals would be *willing to pay* to reduce the risks of death, either through their *revealed preferences*—that is, decisions people actually make in their lives, such as decisions about their work or their retirement plans—or through their *expressed preferences*—that is, what people say in response to hypothetical questions about their preferences. For revealed preferences to be meaningful, individuals must understand the risks in their lives and voluntarily assume those risks—two conditions of autonomous choice that often are not met. For expressed preferences, individuals' answers to hypothetical questions may not accurately indicate how much they would be willing to spend on actual programs to reduce their (and others') risk of death. Individuals' financial situations (including their household income, real estate, and financial solvency) are also likely to have an impact on their expressed willingness to pay.[81]

Even if we rarely put an explicit monetary value on a human life, proponents of CBA often urge such a strategy, notably so in the context of "a statistical life."[82] However, qualitative factors, such as how deaths occur, are often more important to people than purely economic considerations. Moreover, beneficence is often expressed in policies such as rescuing trapped coal miners that symbolize societal benevolence and affirm the value of victims even when these policies would not be supported by a CBA focused on the economic value of life, determined by WTP.

In our judgment, data gained from CBA and other analytic techniques can be made relevant to the formulation and assessment of public policies and can provide valuable information and insights if appropriate qualifications and limits are articulated, but they provide only one set of indicators of appropriate social beneficence. It is often not necessary to put a specific economic value on human life to evaluate different possible risk-reduction policies and to compare their costs. Evaluation may reasonably focus on the lives or the life-years saved, without attempting to convert them into monetary terms. In health care, CBA has now, appropriately, declined in use and importance by comparison to CEA, which often promotes the goal of maximizing QALYs, a topic to which we now turn.[83]

Valuing Quality-Adjusted Life-Years

Quality of life and QALYs. Quality of life is as important as saving lives and years of life in several areas of health policy and health care. Many individuals, when contemplating different treatments for a particular condition, are willing to trade some life-years for improved quality of life during their remaining years. Accordingly, researchers and policymakers have sought measures, called health-adjusted life-years (HALYs), that combine longevity with health status. QALYs are the most widely used type of HALY.[84] The National Institute for Health and Clinical Excellence (NICE), a public body of the Department of Health in the United Kingdom, uses QALYs in evaluations designed for the British system of resource allocation. NICE defines a QALY as "a measure of health outcome which looks at both length of life and quality of life. QALYS are calculated by estimating the years of life remaining for a patient following a particular care pathway and weighting each year with a quality of life score."[85] In short, a QALY is a calculation that takes into account both the quantity and the quality of life produced by medical interventions.

An influential premise underlying use of QALYs is that "if an extra year of healthy (i.e., good quality) life-expectancy is worth one, then an extra year of unhealthy (i.e., poor quality) life-expectancy must be worth less than one (for why otherwise do people seek to be healthy?)."[86] On this scale, the value of the condition of death is zero. Various states of illness or disability better than death but short of full health receive a value between zero and one. Health conditions assessed as worse than death receive a negative value. The value of particular health outcomes depends on the increase in the utility of the health state and the number of years it lasts.[87]

The goal of QALY analysis is to bring length of life and quality of life into a single framework of evaluation.[88] QALYs can be used to monitor the effects of treatments on patients in clinical practice or in clinical trials, to determine what to recommend to patients, to provide information to patients about the effects of different treatments, and to assist in resource allocation in health care. The goal is to make this basis for choices between options as clear and rational as possible.

In an influential case study, British health economist Alan Williams used QALYs to examine the cost-effectiveness of coronary artery bypass graft surgery. In his analysis, bypass grafting compares favorably with pacemakers for heart block. It is superior to heart transplantation and the treatment of end-stage renal failure. He also found that bypass grafting for severe angina and extensive coronary artery disease is more cost-effective than for less severe cases. The rate of survival by itself can be misleading for coronary artery bypass grafting and many other therapeutic procedures that also have a major impact on quality of life. Ultimately, Williams recommended that resources "be redeployed at the margin to procedures for which the benefits to patients are high in relation to the costs."[89]

Nonetheless, the methods for determining quality of life pose many difficulties. Analysts often start with rough measures, such as physical mobility, freedom from pain and distress, and the capacity to perform the activities of daily life and to engage in social interactions. Quality-of-life measures are theoretically attractive as a way to provide information about the ingredients of a good life, but practically difficult to implement. However, some instruments can and should be developed and refined to present meaningful and accurate measures of health-related quality of life. Without such instruments, we are likely to operate with implicit and unexamined views about trade-offs between quantity and quality of life in relation to cost.

Still, these instruments can be misleading because of their built-in ethical assumptions, a problem to which we turn next.

Ethical assumptions of QALYs. Many ethical assumptions are incorporated into QALY-based CEA. Utilitarianism is CEA's philosophical parent, and some of its problems carry over to its offspring, even though there are differences.[90] Implicit in QALY-based CEA is the idea that health maximization is the only relevant objective of health services. But some nonhealth benefits or utilities of health services also contribute to quality of life. As our discussion of silicone gel-filled breast implants noted earlier in this chapter, conditions such as asymmetrical breasts may affect a person's subjective estimate of quality of life and may constitute a source of

distress. The problem is that QALY-based CEAs attach utility only to selected outcomes while neglecting values such as how care is provided (e.g., whether it is personal care) and how it is distributed (e.g., whether universal access is provided).[91]

Related issues arise about whether the use of QALYs in CEA is adequately egalitarian. Proponents of QALY-based CEA hold that each healthy life-year is equally valuable for everyone. A QALY is a QALY, regardless of who possesses it.[92] However, QALY-based CEA may in effect discriminate against older people, because, conditions being equal, saving the life of a younger person is likely to produce more QALYs than saving the life of an older person.[93]

QALY-based CEA also fails to attend adequately to other problems of justice, including the needs of people with disabilities and the needs of the worst off in terms of the severity of their current illness and their health over a lifetime.[94] It does not consider how life-years are distributed among patients, and it may not include efforts to reduce the number of individual victims in its attempts to increase the number of life-years. From this standpoint, no difference exists between saving one person who can be expected to have forty QALYs and saving two people who can be expected to have twenty QALYs each. In principle, CEA will give priority to saving one person with forty expected QALYs over saving two persons with only nineteen expected QALYs each. Hence, QALY-based CEA favors life-years over individual lives, and the number of life-years over the number of individual lives, while failing to recognize that societal and professional obligations of beneficence sometimes require rescuing endangered individual lives.[95]

A tension can easily emerge between QALY-based CEA and the duty to rescue, even though both are ultimately grounded in obligations of beneficence. This tension appeared in a classic effort by the Oregon Health Services Commission to develop a prioritized list of health services so that the state of Oregon could expand its Medicaid coverage to all of its poor citizens. (See our examination of this policy in Chapter 7, pp. 301–2.) A draft priority list ranked some life-saving procedures (e.g., appendectomy for acute appendicitis) below more routine procedures (e.g., capping teeth). About this kind of priority listing, David Hadorn observed: "The cost-effectiveness analysis approach used to create the initial list conflicted directly with the powerful 'Rule of Rescue'—people's perceived duty to save endangered life whenever possible."[96] If unqualified by further ethical considerations, QALY-based CEA's methodological assignment of priority to life-years over individual lives implies that beneficence-based rescue (especially life-saving) is less significant than cost utility, that the distribution of life-years is unimportant, that saving more lives is less important than maximizing the number of life-years, and that quality of life is more important than quantity of life. Each of these priorities needs careful scrutiny in each context in which QALYs are used.

Important questions of justice, fairness, and equity, as well as beneficence challenge both the *conduct* and the *use* of QALY-based CEAs. Some of these challenges can be addressed by modifying underlying assumptions, such as those related to disability and age. However, absent such modifications, it is unclear how far QALY-based CEAs can incorporate relevant concerns of justice, fairness, and equity that reflect social values, beyond individuals' willingness to pay. Equity-weighted CEAs have been proposed that seem attractive,[97] but the combination of QALY and equity in a single CEA is problematic, on grounds of feasibility as well as potential distortion. It seems more reasonable for decision makers to accept QALY-based CEAs, with their assumptions properly examined and modified or corrected, as one major source of input for deliberations. The use of this tentatively accepted input can then be limited and constrained by considerations of justice—a major topic explored further in Chapter 7.

CONCLUSION

In this chapter we have distinguished two principles of beneficence—positive beneficence and utility—and defended the theoretical and practical importance of the distinction between obligatory beneficence and ideal beneficence. We have developed an account of paternalism that makes it possible to justify a restricted range of both soft and hard paternalistic actions. We have nonetheless acknowledged that, in addition to its potential for

disrespect of personal autonomy, a *policy or rule* in law and institutions permitting hard paternalistic actions in professional practice will be dangerous because of the risk of abuse it invites. The fact that physicians are situated to make sound and caring decisions from a position of professional expertise should be one factor, but only one factor, in the on-balance consideration of whether paternalistic interventions in medicine are morally justified.

Finally, we examined formal techniques of analysis—RBA, CBA, and CEA—and concluded that, with suitable qualifications, they are morally unobjectionable ways to explicate the principle of utility, as a principle of beneficence, but that principles of respect for autonomy and justice often should be used to set limits on the uses of these techniques. Chapter 7 develops an account of some principles of justice that began to surface in the final parts of this chapter.

NOTES

1. 1. Bernard Gert presents an aggressive and impressive theory of this sort. He regards beneficence as in the realm of moral ideals, not the realm of moral obligations. See our exegesis and critical evaluation of his theory in Chapter 10, pp. 428–32.
2. 2. W. D. Ross, *The Right and the Good* (Oxford: Clarendon, 1930), p. 21.
3. 3. Peter Singer, "Famine, Affluence, and Morality," *Philosophy & Public Affairs* 1 (1972): 229–43. Richard Arneson generally agrees with Singer but holds that while distance does not change rightness or wrongness of action or inaction, it can, in an act-consequentialist framework, affect an agent's blameworthiness and morally appropriate guilt. See Arneson, "Moral Limits on the Demands of Beneficence?" in *The Ethics of Assistance: Morality and the Distant Needy*, ed. Deen K. Chatterjee (Cambridge: Cambridge University Press, 2004), pp. 33–58.
4. 4. Peter Singer, *Practical Ethics*, 3rd ed. (Cambridge: Cambridge University Press, 2011), chap. 8.
5. 5. Peter Singer, *The Life You Can Save: Acting Now to End World Poverty* (New York: Random House, 2009), especially chaps. 9–10.
6. 6. For assessments of overdemanding theories, see, among others, Liam B. Murphy, "The Demands of Beneficence," *Philosophy & Public Affairs* 22 (1993): 267–92; Murphy, *Moral Demands in Nonideal Theory* (New York: Oxford University Press, 2000); Richard W. Miller, "Beneficence, Duty and Distance," *Philosophy & Public Affairs* 32 (2004): 357–83; Miller, *Globalizing Justice: The Ethics of Poverty and Power* (Oxford: Oxford University Press, 2010); and Brad Hooker, "The Demandingness Objection," in *The Problem of Moral Demandingness*, ed. Timothy Chappell (Basingstoke, UK: Palgrave Macmillan 2009), pp. 148–62.
7. 7. Our formulations are indebted to Eric D'Arcy, *Human Acts: An Essay in Their Moral Evaluation* (Oxford: Clarendon, 1963), pp. 56–57. We added the fourth condition and altered others in his formulation. Our reconstruction profited from Joel Feinberg, *Harm to Others*, vol. 1 of *The Moral Limits of the Criminal Law* (New York: Oxford University Press, 1984), chap. 4.
8. 8. This third condition will need a finer-grained analysis to avoid some problems of what is required if there is a small (but not insignificant) probability of saving millions of lives at minimal cost to a person. It is not plausible to hold that a person has *no obligation* to so act. Condition 3 here could be refined to show that there must be some appropriate proportionality between probability of success, the value of outcome to be achieved, and the sacrifice that the agent would incur. Perhaps the formulation should be "a high ratio of probable benefit relative to the sacrifice made."
9. 9. On the significant role of AIDS activists, see Steven Epstein, *Impure Science: AIDS, Activism, and the Politics of Knowledge* (Berkeley: University of California Press, 1996); and Robert J. Levine, "The Impact of HIV Infection on Society's Perception of Clinical Trials," *Kennedy Institute of Ethics Journal* 4 (1994): 93–98. For some controversies at the time regarding the AIDS activists' goals, see Institute of Medicine (later National Academy of Medicine), *Expanding Access to Investigational Therapies for HIV Infection and AIDS* (Washington, DC: National Academies Press, 1991).
10. 10. US Food and Drug Administration, "Fast Track, Breakthrough Therapy, Accelerated Approval, and Priority Review" (information updated February 23, 2018), available at https://www.fda.gov/forpatients/approvals/fast/ucm20041766.htm (accessed June 9, 2018).

11. 11. Our discussion of these issues is intended to cover a variety of actual and possible expanded access programs. It is not limited to programs that fall under the policies of the US Food and Drug Administration. For the latter, see "Learn about Expanded Access and Other Treatment Options," updated January 4, 2018, available at http://www.fda.gov/ForConsumers/ByAudience/ForPatientAdvocates/AccesstoInvestigationalDrugs/ucm176098.htm (accessed June 7, 2018). In addition, the FDA has a "Parallel Track" policy that "permits wider access to promising new drugs for AIDS/HIV related diseases under a separate 'expanded access' protocol that 'parallels' the controlled clinical trials that are essential to establish the safety and effectiveness of new drugs." See US Food and Drug Administration, "Treatment Use of Investigational Drugs—Information Sheet," available at https://www.fda.gov/RegulatoryInformation/Guidances/ucm126495.htm, as updated March 29, 2018 (accessed June 10, 2018).
12. 12. See Laurie McGinley, "Are Right-to-Try Laws a Last Hope for Dying Patients—or a False Hope?" *Washington Post*, March 26, 2017, available at https://www.washingtonpost.com/national/health-science/are-right-to-try-laws-a-last-hope-for-dying-patients–or-a-cruel-sham/2017/03/26/1aa49c7c-10a2-11e7-ab07-07d9f521f6b5_story.html?utm_term=.061a38dbb205 (accessed June 4, 2018).
13. 13. Lisa Kearns and Alison Bateman-House, "Who Stands to Benefit? Right to Try Law Provisions and Implications," *Therapeutic Innovation & Regulatory Science* 51, no. 2 (2017): 170–76, available at https://med.nyu.edu/pophealth/sites/default/files/pophealth/Kearns%20BatemanHouse%20RTT%20variations%20in%20TIRS.pdf (accessed June 4, 2018); and Elena Fountzilas, Rabih Said, and Apostolia M. Tsimberidou, "Expanded Access to Investigational Drugs: Balancing Patient Safety with Potential Therapeutic Benefits," *Expert Opinion on Investigational Drugs* 27, no. 2 (2018): 155–62, available at https://www.tandfonline.com/doi/full/10.1080/13543784.2018.1430137 (accessed June 4, 2018).
14. 14. Michelle M. Mello and Troyen A. Brennan, "The Controversy over High-Dose Chemotherapy with Autologous Bone Marrow Transplant for Breast Cancer," *Health Affairs* 20 (2001): 101–17; Edward A. Stadtmauer et al., "Conventional-Dose Chemotherapy Compared with High-Dose Chemotherapy Plus Autologous Hematopoietic Stem-Cell Transplantation for Metastatic Breast Cancer," *New England Journal of Medicine* 342 (2000): 1069–76; and Rabiya A. Tuma, "Expanded-Access Programs: Little Heard Views from Industry," *Oncology Times* 30 (August 10, 2008): 19, 22–23. For a thorough review of this history, see Richard A. Rettig, Peter D. Jacobson, Cynthia M. Faquhar, and Wade M. Aubry, *False Hope: Bone Marrow Transplantation for Breast Cancer* (New York: Oxford University Press, 2007).
15. 15. William C. Buhles, "Compassionate Use: A Story of Ethics and Science in the Development of a New Drug," *Perspectives in Biology and Medicine* 54 (2011): 304–15. The case is far more complicated than we report here.
16. 16. Cf. conclusions about post-trial access in National Bioethics Advisory Commission (NBAC), *Ethical and Policy Issues in International Research: Clinical Trials in Developing Countries* (Bethesda, MD: NBAC, April 2001), vol. 1, pp. 64–65, 74, especially Recommendation 4.1, available at https://bioethicsarchive.georgetown.edu/nbac/clinical/Vol1.pdf (accessed August 23, 2018). See also Nuffield Council on Bioethics, *The Ethics of Research Related to Healthcare in Developing Countries* (London: Nuffield Council on Bioethics, 2002), chap. 9, "What Happens Once Research Is Over?" sects. 9.21–31, available at http://nuffieldbioethics.org/wp-content/uploads/2014/07/Ethics-of-research-related-to-healthcare-in-developing-countries-I.pdf (accessed June 7, 2018).
17. 17. David Hume, "Of Suicide," in *Essays Moral, Political, and Literary*, ed. Eugene Miller (Indianapolis, IN: Liberty Classics, 1985), pp. 577–89.
18. 18. See David A. J. Richards, *A Theory of Reasons for Action* (Oxford: Clarendon, 1971), p. 186; Allen Buchanan, "Justice as Reciprocity vs. Subject-Centered Justice," *Philosophy & Public Affairs* 19 (1990): 227–52; Lawrence Becker, *Reciprocity* (Chicago: University of Chicago Press, 1990); and Aristotle, *Nicomachean Ethics*, bks. 8–9.
19. 19. See William F. May, "Code and Covenant or Philanthropy and Contract?" in *Ethics in Medicine*, ed. Stanley Reiser, Arthur Dyck, and William Curran (Cambridge, MA: MIT Press, 1977), pp. 65–76; and May, *The Healer's Covenant: Images of the Healer in Medical Ethics*, 2nd ed. (Louisville, KY: Westminster-John Knox Press, 2000).
20. 20. Institute of Medicine (later National Academy of Medicine) of the National Academies, Roundtable on Evidence-Based Medicine, *The Learning Healthcare System: Workshop Summary*, ed. LeighAnne Olsen, Dara Aisner, and J. Michael McGinnis (Washington, DC: National Academies Press, 2007), esp.

chap. 3, available at http://www.nap.edu/catalog/11903.html (accessed June 7, 2018); Ruth R. Faden, Nancy E. Kass, Steven N. Goodman, Peter Pronovost, Sean Tunis, and Tom L. Beauchamp, "An Ethics Framework for a Learning Healthcare System," *Hastings Center Report* (Special Report) 43 (2013): S16–S27; and Committee on the Learning Health Care System in America, Institute of Medicine (now National Academy of Medicine) of the National Academies, *Best Care at Lower Cost: The Path to Continuously Learning Health Care in America*, ed. Mark Smith, Robert Saunders, Leigh Stuckhardt, and J. Michael McGinnis (Washington, DC: National Academies Press, 2013), available at https://www.nap.edu/read/13444/chapter/1 (accessed June 25, 2018).

21. 21. For an ethical evaluation of Israel's policy, see Jacob Lavee and Dan W. Brock, "Prioritizing Registered Donors in Organ Allocation: An Ethical Appraisal of the Israeli Organ Transplant Law," *Current Opinion in Critical Care* 18, no. 6 (2012): 707–11. They assess the law to be basically sound but in need of modification (especially priority for first-degree relatives). A defense of prioritizing registered donors in allocation appears in Gil Siegal and Richard Bonnie, "Closing the Organ Donation Gap: A Reciprocity-Based Social Contract Approach," *Journal of Law, Medicine & Ethics* 34 (2006): 415–23. For an analysis and assessment of the two models we have identified, see James F. Childress and Catharyn T. Liverman, eds., *Organ Donation: Opportunities for Action* (Washington, DC: National Academies Press, 2006), pp. 253–59, which argues against both models "because of insuperable practical problems in implementing them fairly" (p. 253).
22. 22. *Epidemics*, 1:11, in *Hippocrates*, vol. 1, ed. W. H. S. Jones (Cambridge, MA: Harvard University Press, 1923), p. 165.
23. 23. See Tom L. Beauchamp, "The Concept of Paternalism in Biomedical Ethics," *Jahrbuch für Wissenschaft und Ethik* 14 (2010): 77–92, which presents the following alternative definition: "Paternalism is the intentional overriding of one person's autonomous choices or actions by another person, where the person who overrides justifies the action by appeal to the goal of benefiting or of preventing or mitigating harm to the person whose choices or actions are overridden." Under this definition, a person's choices or actions must be substantially autonomous for an intervention to qualify as paternalistic.
24. 24. See Donald VanDeVeer, *Paternalistic Intervention: The Moral Bounds on Benevolence* (Princeton, NJ: Princeton University Press, 1986), pp. 16–40; John Kleinig, *Paternalism* (Totowa, NJ: Rowman & Allanheld, 1983), pp. 6–14; and James F. Childress, *Who Should Decide? Paternalism in Health Care* (New York: Oxford University Press, 1982). See also Childress, "Paternalism and Autonomy in Medical Decision-Making," in *Frontiers in Medical Ethics: Applications in a Medical Setting*, ed. Virginia Abernethy (Cambridge, MA: Ballinger, 1980), pp. 27–41; and Childress, "Paternalism in Health Care and Public Policy," in *Principles of Health Care Ethics*, 2nd ed., ed., Richard E. Ashcroft, Angus Dawson, Heather Draper, and John McMillan (Chichester, UK: John Wiley, 2007), pp. 223–31.
25. 25. This case is formulated on the basis of, and incorporates language from, Margaret A. Drickamer and Mark S. Lachs, "Should Patients with Alzheimer's Be Told Their Diagnosis?" *New England Journal of Medicine* 326 (April 2, 1992): 947–51. For diagnostic guidelines for Alzheimer's disease (updated in January 2011), see the information provided by the National Institute of Aging at https://www.nia.nih.gov/health/alzheimers-disease-diagnostic-guidelines (accessed June 7, 2018). Only an autopsy after the patient's death can provide a definitive diagnosis of Alzheimer's disease.
26. 26. First introduced as the distinction between strong and weak paternalism by Joel Feinberg, "Legal Paternalism," *Canadian Journal of Philosophy* 1 (1971): 105–24, esp. pp. 113, 116. See, further, Feinberg, *Harm to Self*, vol. 3 of *The Moral Limits of the Criminal Law* (New York: Oxford University Press, 1986), esp. pp. 12ff.
27. 27. See Cass R. Sunstein and Richard H. Thaler, "Libertarian Paternalism Is Not an Oxymoron," *University of Chicago Law Review* 70 (Fall 2003): 1159–202; Thaler and Sunstein, *Nudge: Improving Decisions about Health, Wealth, and Happiness* (New Haven, CT: Yale University Press, 2008); and Sunstein, *Why Nudge? The Politics of Libertarian Paternalism* (New Haven, CT: Yale University Press, 2014).
28. 28. Erich H. Loewy, "In Defense of Paternalism," *Theoretical Medicine and Bioethics* 26 (2005): 445–68.
29. 29. Childress, *Who Should Decide? Paternalism in Health Care*, p. 18.
30. 30. Sunstein and Thaler, "Libertarian Paternalism Is Not an Oxymoron," p. 1159. See also Thaler and Sunstein, "Libertarian Paternalism," *American Economics Review* 93 (2003): 175–79.

31. 31. Christine Jolls and Cass R. Sunstein, "Debiasing through Law," *Journal of Legal Studies* 33 (January 2006): 232.
32. 32. See Edward L. Glaeser, "Symposium: Homo Economicus, Homo Myopicus, and the Law and Economics of Consumer Choice: Paternalism and Autonomy," *University of Chicago Law Review* 73 (Winter 2006): 133–57. The work of Thaler and Sunstein has spawned a large body of both critical and supportive literature. For a libertarian critique of their views on libertarian paternalism, see Richard A. Epstein, "Libertarian Paternalism Is a Nice Phrase for Controlling People," *Federalist*, 2018, available at http://thefederalist.com/2018/04/26/libertarian-paternalism-nice-phrase-controlling-people/ (accessed August 18, 2018). For criticisms of soft as well as hard paternalism, see Christopher Snowdon, *Killjoys: A Critique of Paternalism* (London: Institute of Economic Affairs, 2017); Mark D. White, *The Manipulation of Choice: Ethics and Libertarian Paternalism* (London: Palgrave Macmillan, 2013), which argues "vehemently" against libertarian paternalism and nudges; and Sherzod Abdukadirov, ed., *Nudge Theory in Action: Behavioral Design in Policy and Markets* (London: Palgrave Macmillan, 2016), which includes several critical essays. Proponents include, in addition to the literature cited in other notes, Sigal R. Ben-Porath, *Tough Choices: Structured Paternalism and the Landscape of Choice* (Princeton, NJ: Princeton University Press, 2010); and Sarah Conly, *Against Autonomy: Justifying Coercive Paternalism* (Cambridge: Cambridge University Press, 2013). Some collections of essays include both critics and defenders: See Christian Coons and Michael Weber, eds., *Paternalism: Theory and Practice* (Cambridge: Cambridge University Press, 2013); and I. Glenn Cohen, Holly Fernandez Lynch, and Christopher T. Robertson, eds., *Nudging Health: Health Law and Behavioral Economics* (Baltimore, MD: Johns Hopkins University Press, 2016).
33. 33. Ronald Bayer and Jennifer Stuber, "Tobacco Control, Stigma, and Public Health: Rethinking the Relations," *American Journal of Public Health* 96 (January 2006): 47–50; and Glaeser, "Symposium: Homo Economicus, Homo Myopicus, and the Law and Economics of Consumer Choice," pp. 152–53. Stigmatization has emerged in efforts to reduce obesity, opioid abuse, and other harmful behaviors. For a recognition of the legitimate role, within limits, of stigmatization in public health, see A. Courtwright, "Stigmatization and Public Health Ethics," *Bioethics* 27 (2013): 74–80; and Daniel Callahan, "Obesity: Chasing an Elusive Epidemic," *Hastings Center Report* 43, no. 1 (January–February 2013): 34–40. For a rejection of stigmatization in campaigns against obesity because of its several negative impacts, see C. J. Pausé, "Borderline: The Ethics of Fat Stigma in Public Health," *Journal of Law, Medicine & Ethics* 45 (2017): 510–17.
34. 34. Bayer and Stuber, "Tobacco Control, Stigma, and Public Health: Rethinking the Relations," p. 49.
35. 35. W. Kip Vicusi, "The New Cigarette Paternalism," *Regulation* (Winter 2002–3): 58–64.
36. 36. For interpretations of (hard) paternalism as insult, disrespect, and treatment of individuals as unequals, see Ronald Dworkin, *Taking Rights Seriously* (Cambridge, MA: Harvard University Press, 1978), pp. 262–63; and Childress, *Who Should Decide?* chap. 3.
37. 37. Gerald Dworkin, "Paternalism," *Monist* 56 (1972): 65. See also Gerald Dworkin, "Paternalism," in *The Stanford Encyclopedia of Philosophy* (Winter 2017 Edition), ed. Edward N. Zalta, available at https://plato.stanford.edu/archives/win2017/entries/paternalism/ (accessed June 9, 2018).
38. 38. See Gerald Dworkin, "Paternalism," *Monist* 56 (1972); and John Rawls, *A Theory of Justice* (Cambridge, MA: Harvard University Press, 1971; rev. ed., 1999), pp. 209, 248–49 (1999: pp. 183–84, 218–20).
39. 39. Gerald Dworkin says, "The reasons which support paternalism are those which support any altruistic action—the welfare of another person." "Paternalism," in *Encyclopedia of Ethics*, ed. Lawrence Becker (New York: Garland, 1992), p. 940. For a variety of consent and nonconsent defenses of paternalism, see Kleinig, *Paternalism*, pp. 38–73; and John Kultgen, *Autonomy and Intervention: Paternalism in the Caring Life*: (New York: Oxford University Press, 1995), esp. chaps. 9, 11, 15.
40. 40. We take a constrained-balancing approach to the conflict between respect for autonomy and beneficence to a particular person. Another approach could develop a specification of beneficence and respect for autonomy that would rule out all hard paternalistic interventions. The specification could take the following form: "When a person's actions are substantially autonomous and create the risk of harm to himself or herself, without imposing significant harms or burdens on others or the society, we should not act paternalistically beyond the use of modest means such as persuasion." Determining whether such a

specification could be rendered coherent with our overall approach would require more attention than we can devote here.

41. 41. See further our discussion of staged disclosure of information in Chapter 8, pp. 330–34.
42. 42. Deborah M. Stone, Thomas R. Simon, Katherine A. Fowler, et al., "Vital Signs: Trends in State Suicide Rates—United States, 1999–2016 and Circumstances Contributing to Suicide—27 States, 2015," *Morbidity and Mortality Weekly Report* 67 (2018): 617–24, available at http://dx.doi.org/10.15585/mmwr.mm6722a1 (accessed June 6, 2018).
43. 43. We do not here address philosophical problems surrounding the definition of suicide. On this matter, see Tom L. Beauchamp, "Suicide," in *Matters of Life and Death*, 3rd ed., ed. Tom Regan (New York: Random House, 1993), esp. part 1; John Donnelly, ed., *Suicide: Right or Wrong?* (Buffalo, NY: Prometheus Books, 1991), part 1; and Michael Cholbi, *Suicide: The Philosophical Dimensions* (Toronto: Broadview Press, 2011), chap. 1. In Chapter 5 we examined reasons for not labeling physician-assisted death, in which the patient performs the final act, as physician-assisted "suicide."
44. 44. See James Rachels, "Barney Clark's Key," *Hastings Center Report* 13 (April 1983): 17–19, esp. 17.
45. 45. This case is presented in Marc Basson, ed., *Rights and Responsibilities in Modern Medicine* (New York: Alan R. Liss, 1981), pp. 183–84.
46. 46. Glanville Williams, "Euthanasia," *Medico-Legal Journal* 41 (1973): 27.
47. 47. See President's Commission for the Study of Ethical Problems in Medicine and Biomedical and Behavioral Research, *Deciding to Forego Life-Sustaining Treatment: Ethical, Medical, and Legal Issues in Treatment Decisions* (Washington, DC: US Government Printing Office, March 1983), p. 37.
48. 48. Betty Rollin, *Last Wish* (New York: Linden Press Simon & Schuster, 1985).
49. 49. Childress, *Who Should Decide?* chap. 1. See also Timothy E. Quill and Howard Brody, "Physician Recommendations and Patient Autonomy: Finding a Balance between Physician Power and Patient Choice," *Annals of Internal Medicine* 125 (1996): 763–69; Allan S. Brett and Laurence B. McCullough, "When Patients Request Specific Interventions: Defining the Limits of the Physician's Obligation," *New England Journal of Medicine* 315 (November 20, 1986): 1347–51; and Brett and McCullough, "Addressing Requests by Patients for Nonbeneficial Interventions," *JAMA: Journal of the American Medical Association* 307 (January 11, 2012): 149–50.
50. 50. We have adapted this case from "The Refusal to Sterilize: A Paternalistic Decision," in *Rights and Responsibilities in Modern Medicine*, ed. Basson, pp. 135–36.
51. 51. See Steven H. Miles, "Informed Demand for Non-Beneficial Medical Treatment," *New England Journal of Medicine* 325 (August 15, 1991): 512–15; and Ronald E. Cranford, "Helga Wanglie's Ventilator," *Hastings Center Report* 21 (July–August 1991): 23–24.
52. 52. Catherine A. Marco and Gregory L. Larkin, "Case Studies in 'Futility'—Challenges for Academic Emergency Medicine," *Academic Emergency Medicine* 7 (2000): 1147–51.
53. 53. See further Lawrence J. Schneiderman, Nancy S. Jecker, and Albert R. Jonsen, "The Abuse of Futility," *Perspectives in Biology and Medicine* 60 (2017): 295–313. For a rich international exploration of concepts of and practices related to medical futility, see Alireza Bagheri, ed., *Medical Futility: A Cross-National Study* (London: Imperial College Press, 2013).
54. 54. For a helpful introduction to risk, see Baruch Fischhoff and John Kadvany, *Risk: A Very Short Introduction* (Oxford: Oxford University Press, 2011).
55. 55. See, for example, Charles Yoe, *Primer on Risk Analysis: Decision Making under Uncertainty* (Boca Raton, FL: CRC Press, 2012). A fuller discussion appears in Yoe, *Principles of Risk Analysis: Decision Making under Uncertainty* (Boca Raton, FL: CRC Press, 2012).
56. 56. See Sheila Jasanoff, "Acceptable Evidence in a Pluralistic Society," in *Acceptable Evidence: Science and Values in Risk Management*, ed. Deborah G. Mayo and Rachelle D. Hollander (New York: Oxford University Press, 1991).
57. 57. See Richard Wilson and E. A. C. Crouch, "Risk Assessment and Comparisons: An Introduction," *Science* 236 (April 17, 1987): 267–70; Wilson and Crouch, *Risk-Benefit Analysis* (Cambridge, MA: Harvard University Center for Risk Analysis, 2001); and Baruch Fischoff, "The Realities of Risk-Cost-Benefit Analysis," *Science* 350 (6260) (October 2015): 527, aaa6516-aaa651, available at https://www.researchgate.net/publication/283330070_The_realities_of_risk-cost-benefit_analysis (accessed July 14, 2018).

58. 58. For a summary of this dispute—and a strongly stated argument in favor of the American Society position—see Brian J. Morris, Jeffrey D. Klausner, John N. Krieger, et al., "Canadian Pediatrics Society Position Statement on Newborn Circumcision: A Risk-Benefit Analysis Revisited," *Canadian Journal of Urology* 23, no. 5 (October 2016): 8495–502. This study performs its own risk-benefit analysis and claims to be "more inclusive" than the Canadian Society study of 2015, which is assessed as "at odds with the evidence" and as suffering from serious "errors in its risk-benefit analysis." Of the six authors of this study, two are from Canada and three from the United States. The first author is from Australia.
59. 59. Curt D. Burberg, Arthur A. Levin, Peter A. Gross, et al., "The FDA and Drug Safety," *Archives of Internal Medicine* 166 (October 9, 2006): 1938–42; and Alina Baciu, Kathleen Stratton, and Sheila P. Burke, eds., *The Future of Drug Safety: Promoting and Protecting the Health of the Public* (Washington, DC: National Academies Press, 2006).
60. 60. David A. Kessler, "Special Report: The Basis of the FDA's Decision on Breast Implants," *New England Journal of Medicine* 326 (June 18, 1992): 1713–15. All references to Kessler's views are to this article.
61. 61. See Marcia Angell, "Breast Implants—Protection or Paternalism?" *New England Journal of Medicine* 326 (June 18, 1992): 1695–96. Angell's criticisms also appear in her *Science on Trial: The Clash of Medical Evidence and the Law in the Breast Implant Case* (New York: Norton, 1996). See also Jack C. Fisher, *Silicone on Trial: Breast Implants and the Politics of Risk* (New York: Sager Group LLC, 2015), which is sharply critical of the FDA's early decisions.
62. 62. For reviews and evaluations of the scientific data, see E. C. Janowsky, L. L. Kupper, and B. S. Hulka, "Meta-Analyses of the Relation between Silicone Breast Implants and the Risk of Connective Tissue Diseases," *New England Journal of Medicine* 342 (2000): 781–90; *Silicone Gel Breast Implants: Report of the Independent Review Group* (Cambridge, MA: Jill Rogers Associates, 1998); and S. Bondurant, V. Ernster, and R. Herdman, eds., *Safety of Silicone Breast Implants* (Washington, DC: National Academies Press, 2000).
63. 63. "FDA Approves Silicone Gel-Filled Breast Implants after In-Depth Evaluation," *FDA News*, November 17, 2006. Since that time, the FDA has approved five silicone gel-filled breast implants. See US Food and Drug Administration, *Silicone Gel-Filled Breast Implants* (with several links), updated March 26, 2018, available at https://www.fda.gov/MedicalDevices/ProductsandMedicalProcedures/ImplantsandProsthetics/BreastImplants/ucm063871.htm (accessed June 4, 2018).
64. 64. Center for Devices and Radiological Health, US Food and Drug Administration, *FDA Update on the Safety of Silicone Gel-Filled Breast Implants* (June 2011), available at http://www.fda.gov/downloads/MedicalDevices/ProductsandMedicalProcedures/ImplantsandProsthetics/BreastImplants/UCM260090.pdf (accessed June 4, 2018). Further controversy erupted in late 2018, when a study of long-term outcomes in close to 100,000 women with breast implants found an association between the implants and four health problems (melanoma and three auto-immune disorders). See Christopher J. Coroneos, Jesse C. Selber, Anaeze C. Offodile et al., "US FDA Breast Implant Postapproval Studies: Long-term Outcomes in 99,993 Patients," *Annals of Surgery* 269, no. 1 (January 2019). See also Binita S. Ashar, "Assessing the Risks of Breast Implants and FDA's Vision for the National Breast Implant Registry," *Annals of Surgery* 269, no. 1 (January 2019). While noting the study's methodological limitations, the FDA decided to convene a public meeting of its Medical Devices Advisory Committee to address the issues. After this meeting in March 2019, the FDA decided not to ban any breast implants but to ensure that more information about risks is available to prospective users, including about the increased risk of breast implant-associated anaplastic large cell lymphoma, especially in users of textured implants. *Statement from FDA Principal Deputy Commissioner Amy Abernethy, M.D., Ph.D., and Jeff Shuren, M.D., J.D., director of the FDA's Center for Devices and Radiological Health on FDA's new efforts to protect women's health and help to ensure the safety of breast implants,* May 02, 2019. Available at https://www.fda.gov/news-events/press-announcements/statement-fda-principal-deputy-commissioner-amy-abernethy-md-phd-and-jeff-shuren-md-jd-director-fdas (accessed May 15, 2019).
65. 65. National Academies of Sciences, Engineering, and Medicine, *Pain Management and the Opioid Epidemic: Balancing Societal and Individual Benefits and Risks of Prescription Opioid Use* (Washington, DC: National Academies Press, 2017). Our paragraphs on this subject draw heavily on this report. See

also National Institute on Drug Abuse, "Opioid Overdose Crisis," as revised March 2018, available at https://www.drugabuse.gov/drugs-abuse/opioids/opioid-overdose-crisis (accessed July 14, 2018); and Owen Amos, "Why Opioids Are Such an American Problem," *BBC News*, Washington DC, October 25, 2017, available at https://www.bbc.com/news/world-us-canada-41701718 (accessed July 14, 2018).

66. 66. National Academies of Sciences, Engineering, and Medicine, *Pain Management and the Opioid Epidemic*, esp. chap. 6.
67. 67. See Paul Slovic, "Perception of Risk," *Science* 236 (April 17, 1987): 280–85; and Slovic, *The Perception of Risk* (London: Earthscan, 2000).
68. 68. See Cass Sunstein, *Laws of Fear: Beyond the Precautionary Principle* (Cambridge: Cambridge University Press, 2005) and his *Risk and Reason* (Cambridge: Cambridge University Press, 2002).
69. 69. For defenses of the precautionary principle, see United Nations Educational, Scientific and Cultural Organization (UNESCO), *The Precautionary Principle* (2005), available at http://unesdoc.unesco.org/images/0013/001395/139578e.pdf (accessed June 4, 2018); Poul Harremoës, David Gee, Malcolm MacGarvin, et al., *The Precautionary Principle in the 20th Century: Late Lessons from Early Warnings* (London: Earthscan, 2002); Tim O'Riordan, James Cameron, and Andrew Jordan, eds., *Reinterpreting the Precautionary Principle* (London: Earthscan, 2001); Carl Cranor, "Toward Understanding Aspects of the Precautionary Principle," *Journal of Medicine and Philosophy* 29 (June 2004): 259–79; and Elizabeth Fisher, Judith Jones, and René von Schomberg, eds., *Implementing the Precautionary Principle: Perspectives and Prospects* (Northampton, MA: Edward Elgar, 2006). For critical perspectives on the precautionary principle, see Sunstein, *Laws of Fear: Beyond the Precautionary Principle*; H. Tristram Engelhardt, Jr., and Fabrice Jotterand, "The Precautionary Principle: A Dialectical Reconsideration," *Journal of Medicine and Philosophy* 29 (June 2004): 301–12; and Russell Powell, "What's the Harm? An Evolutionary Theoretical Critique of the Precautionary Principle," *Kennedy Institute of Ethics Journal* 20 (2010): 181–206.
70. 70. See P. Sandin, "Dimensions of the Precautionary Principle," *Human and Ecological Risk Assessment* 5 (1999): 889–907.
71. 71. Sunstein, *Laws of Fear: Beyond the Precautionary Principle.* See also Engelhardt and Jotterand, "The Precautionary Principle: A Dialectical Reconsideration"; and Søren Holm and John Harris, "Precautionary Principle Stifles Discovery" (correspondence), *Nature* 400 (July 1999): 398.
72. 72. See Christian Munthe, *The Price of Precaution and the Ethics of Risk* (New York: Springer, 2011).
73. 73. Lauren Hartzell-Nichols proposes a variety of precautionary principles rather than *the* precautionary principle. See "From 'the' Precautionary Principle to Precautionary Principles," *Ethics, Policy & Environment* 16 (2013): 308–20. We find the language of precautionary approach or precautionary process more fruitful.
74. 74. Cf. Cass Sunstein, *Laws of Fear: Beyond the Precautionary Principle;* and Richard A. Posner, *Catastrophe: Risk and Response* (New York: Oxford University Press, 2004).
75. 75. Christian Munthe, *The Price of Precaution and the Ethics of Risk*, p. 164. See also the review of this book by Lauren Hartzell-Nichols, "The Price of Precaution and the Ethics of Risk," *Ethics, Policy & Environment* 17 (2014): 116–18. While holding onto the precautionary principle as a normative principle for disproportionate threats (in contrast to manageable risks), Alan Randall incorporates the principle into a responsible strategy of risk management. See Randall, *Risk and Precaution* (Cambridge: Cambridge University Press, 2011), esp. chaps. 12 and 13.
76. 76. See several chapters in O'Riordan, Cameron, and Jordan, eds., *Reinterpreting the Precautionary Principle.* See also the recent debate about blood transfusion risks in the *American Journal of Bioethics (AJOB)* 17, no. 3 (March, 2017): 32–59. The target article by Koen Kramer, Hans L. Zaaijer, and Marcel F. Verweij, "The Precautionary Principle and the Tolerability of Blood Transfusion Risks" (pp. 32–43) develops three constraints on any precautionary principle: consistency, avoiding counterproductivity, and proportionality. Among the several responses, Anthony Vernillo's "The Precautionary Petard: Who Should Tolerate Blood Transfusion Risks?" (pp. 54–55) takes a perspective similar to ours.
77. 77. See Jonathan Zander, *The Application of the Precautionary Principle in Practice: Comparative Dimensions* (New York: Cambridge University Press, 2010), which notes variations in the application of the precautionary principle in Europe, as well as in the United States. *The Reality of Precaution: Comparing Risk Regulation in the United States and Europe*, ed. Jonathan B. Wiener et al. (New York: Routledge, 2010) also challenges the claim that Europe is more precautionary than the United States.

78. 78. Some examinations of analytic methods include cost-utility analysis (CUA), as distinguished from CEA, while other discussions, particularly in the United States, treat CUA as a variant of CEA, as we do. Both of the following books follow the latter convention: Michael F. Drummond, Mark J. Sculpher, Karl Claxton, Greg L. Stoddart, and George W. Torrance, *Methods for the Economic Evaluation of Health Care Programmes*, 4th ed. (New York: Oxford University Press, 2015); and Peter J. Neumann, Gillian D. Sanders, Louise B. Russell, Joanna E. Siegel, and Theodore G. Ganiats, eds., *Cost-Effectiveness in Health and Medicine*, 2nd ed. (New York: Oxford University Press, 2017).
79. 79. Our description of these analytic techniques draws on Neumann, Sanders, Russell, Siegel, and Ganiats, eds., *Cost-Effectiveness in Health and Medicine*, 2nd ed. (and the earlier edition); and Wilhelmine Miller, Lisa A. Robinson, and Robert S. Lawrence, eds., *Valuing Health for Regulatory Effectiveness Analysis* (Washington, DC: National Academies Press, 2006). See also Peter J. Neumann, *Using Cost-Effectiveness Analysis to Improve Health Care: Opportunities and Barriers* (New York: Oxford University Press, 2005). Attention to comparative effectiveness analysis in the United States appears, in part, to be an attempt to avoid facing the trade-offs between costs, on the one hand, and effectiveness and benefits, on the other. See Uwe E. Reinhardt, "'Cost-Effectiveness Analysis' and U.S. Health Care," *New York Times*, March 13, 2009, available at https://economix.blogs.nytimes.com/2009/03/13/cost-effectiveness-analysis-and-us-health-care/ (accessed July 14, 2018).
80. 80. On this now classic example, see Duncan Neuhauser and Ann M. Lewicki, "What Do We Gain from the Sixth Stool Guaiac?" *New England Journal of Medicine* 293 (July 31, 1975): 226–28. See also "American Cancer Society Report on the Cancer-Related Checkup," *CA—A Cancer Journal for Clinicians* 30 (1980): 193–240, which recommended the full set of six guaiac tests.
81. 81. See Emma McIntosh et al., "Applied Cost-Benefit Analysis in Health Care: An Empirical Application in Spinal Surgery," in *Applied Methods of Cost-Benefit Analysis in Health Care*, ed. Emma McIntosh et al. (New York: Oxford University Press, 2010), pp. 139–57, esp. 153–54, which focuses on one type of willingness to pay in the context of spinal surgery.
82. 82. For an illuminating analysis and defense of valuing life (or valuing statistical life), properly qualified, in public regulation, see Cass R. Sunstein, *Valuing Life: Humanizing the Regulatory State* (Chicago: University of Chicago Press, 2014), esp. chaps. 4 and 5.
83. 83. For a philosophical critique of CBA, see Elizabeth Anderson, *Values in Ethics and Economics* (Cambridge, MA: Harvard University Press, 1993), esp. chap. 9; Matthew D. Adler, *Well-Being and Fair Distribution: Beyond Cost-Benefit Analysis* (New York: Oxford University Press, 2012), esp. pp. 88–114; and Peter A. Ubel, *Pricing Life: Why It's Time for Health Care Rationing* (Cambridge, MA: MIT Press, 2000), esp. p. 68.
84. 84. See Miller, Robinson, and Lawrence, eds., *Valuing Health for Regulatory Cost-Effectiveness Analysis.* For an examination and a call for further clarification of different types of measures, see Marthe R. Gold, David Stevenson, and Dennis G. Fryback, "HALYs and QALYs and DALYs, Oh My: Similarities and Differences in Summary Measures of Population Health," *Annual Review of Public Health* 23 (2002): 115–34. For a critical examination of disability-adjusted life years (DALYs), see Sudhir Anand and Kara Hanson, "Disability-Adjusted Life Years: A Critical Review," in *Public Health, Ethics, and Equity*, ed. Sudhir Anand, Fabienne Peter, and Amartya Sen (Oxford: Oxford University Press, 2004), chap. 9.
85. 85. National Institute for Health and Clinical Excellence, *Social Value Judgements: Principles for the Development of NICE Guidance*, 2nd ed. (2008), p. 35, available at https://www.nice.org.uk/media/default/about/what-we-do/research-and-development/social-value-judgements-principles-for-the-development-of-nice-guidance.pdf (accessed June 7, 2018).
86. 86. Alan Williams, "The Importance of Quality of Life in Policy Decisions," in *Quality of Life: Assessment and Application*, ed. Stuart R. Walker and Rachel M. Rosser (Boston: MTP Press, 1988), p. 285.
87. 87. See Erik Nord, *Cost-Value Analysis in Health Care: Making Sense out of QALYs* (Cambridge: Cambridge University Press, 1999), passim; and Neumann, Sanders, Russell, Siegel, and Ganiats, eds., *Cost-Effectiveness in Health and Medicine*, 2nd ed., passim.
88. 88. See David Eddy, "Cost-Effectiveness Analysis: Is It Up to the Task?" *Journal of the American Medical Association* 267 (June 24, 1992): 3344. On "conventional" QALYs, see Milton C. Weinstein, George Torrance, and Alastair McGuire, "QALYs: The Basics," *Value in Health* 12, Supplement 1 (2009): S5–S9.

89. 89. Alan Williams, "Economics of Coronary Artery Bypass Grafting," *British Medical Journal* 291 (August 3, 1985): 326–29. See also M. C. Weinstein and W. B. Stason, "Cost-Effectiveness of Coronary Artery Bypass Surgery," *Circulation* 66, Suppl. 5, pt. 2 (1982): III, 56–66. On the nature of and guidelines for coronary artery bypass graft surgery, see Kim A. Eagle, Robert A. Guyton, Ravin Davidoff, et al., "ACC/AHA 2004 Guideline Update for Coronary Artery Bypass Graft Surgery: A Report of the American College of Cardiology/American Heart Association Task Force on Practice Guidelines," *Circulation* 110 (2004), available at http://circ.ahajournals.org/content/circulationaha/110/14/e340.full.pdf (accessed July 13, 2018).
90. 90. See Paul Menzel, Marthe R. Gold, Erik Nord, et al., "Toward a Broader View of Values in Cost-Effectiveness Analysis of Health," *Hastings Center Report* 29 (May–June 1999): 7–15. For a defense of the utilitarian perspective of CEA and QALYs, see John McKie, Jeff Richardson, and Helga Kuhse, *The Allocation of Health Care Resources: An Ethical Evaluation of the 'QALY' Approach* (Aldershot, England: Ashgate, 1998). See also Joshua Cohen, "Preferences, Needs and QALYs," *Journal of Medical Ethics* 22 (1996): 267–72; Dan W. Brock, "Ethical Issues in the Use of Cost Effectiveness Analysis for the Prioritisation of Health Care Resources," in *Public Health, Ethics, and Equity*, ed. Peter Anand and Amartya Sen, chap. 10; Madison Powers and Ruth Faden, *Social Justice: The Moral Foundations of Public Health and Health Policy* (New York: Oxford University Press, 2006), chap. 6; and Powers and Faden, *Structural Injustice: Power, Advantage, and Human Rights* (New York: Oxford University Press, 2019).
91. 91. Gavin Mooney, "QALYs: Are They Enough? A Health Economist's Perspective," *Journal of Medical Ethics* 15 (1989): 148–52.
92. 92. Alan Williams, "The Importance of Quality of Life in Policy Decisions," in *Quality of Life*, ed. Walker and Rosser, p. 286; and Williams, "Economics, QALYs and Medical Ethics—A Health Economist's Perspective," *Health Care Analysis* 3 (1995): 221–26.
93. 93. Some proposals to modify or limit QALY-based CEA by societal values would require even a lower weighting for the elderly, in line with dominant critical social values. See, for example, Nord, *Cost-Value Analysis in Health Care*; Menzel et al., "Toward a Broader View of Values in Cost-Effectiveness Analysis of Health"; and Ubel, *Pricing Life.*
94. 94. For these and other equity concerns, see Dan W. Brock, Norman Daniels, Peter J. Neumann, and Joanna E. Siegel, "Ethical and Distributive Considerations," in *Cost-Effectiveness in Health and Medicine*, 2nd ed., ed. Neumann, Sanders, Russell, Siegel, and Ganiats, pp. 319–341; Erik Nord, Norman Daniels, and Mark Kamlet, "QALYs: Some Challenges," *Value in Health* 12, Supplement 1 (2009): S10–15; Erik Nord, "Some Ethical Corrections to Valuing Health Programs in Terms of Quality-Adjusted Life Years (QALYs)," *AMA Journal of Ethics*, Virtual Mentor, 7, no. 2 (February 2005). For an effort to calibrate QALYs to ensure the equality of persons with disabilities, see Donald Franklin, "Calibrating QALYs to Respect Equality of Persons," *Utilitas* 29, no. 1 (March 2017): 65–87.
95. 95. See the careful and qualified framework surrounding the use of QALYs by NICE, *Social Value Judgements*, sections 3–4, 7–8, esp. 4.2. John Harris argues that QALYs are a "life-threatening device," because they suggest that life-years rather than individual lives are valuable. "QALYfying the Value of Life," *Journal of Medical Ethics* 13 (1987): 117–23. See also Peter Singer, John McKie, Helga Kuhse, and Jeff Richardson, "Double Jeopardy and the Use of QALYs in Health Care Allocation," *Journal of Medical Ethics* 21 (1995): 144–50; John Harris, "Double Jeopardy and the Veil of Ignorance—A Reply," *Journal of Medical Ethics* 21 (1995): 151–57; John McKie, Helga Kuhse, Jeff Richardson, and Peter Singer, "Double Jeopardy, the Equal Value of Lives and the Veil of Ignorance: A Rejoinder to Harris," *Journal of Medical Ethics* 22 (1996): 204–8.
96. 96. David C. Hadorn, "Setting Health Care Priorities in Oregon: Cost-Effectiveness Meets the Rule of Rescue," *Journal of the American Medical Association* 265 (May 1, 1991): 2218; and David C. Hadorn, "The Oregon Priority-setting Exercise: Cost-Effectiveness and the Rule of Rescue, Revisited," *Medical Decision Making* 16 (1996): 117–19. See, further, Peter Ubel, George Loewenstein, Dennis Scanlon, and Mark Kamlet, "Individual Utilities Are Inconsistent with Rationing Choices: A Partial Explanation of Why Oregon's Cost-Effectiveness List Failed," *Medical Decision Making* 16 (1996): 108–16; and John McKie and Jeff Richardson, "The Rule of Rescue," *Social Science & Medicine* 56 (2003): 2407–19. We return to the Oregon experiment in Chapter 7, pp. 301–2.

97. 97. See Erik Nord, “Cost-Value Analysis of Health Interventions: Introduction and Update on Methods and Preference Data,” *PharmacoEconomics* 33 (2015): 89–95; and Brock, Daniels, Neumann, and Siegel, “Ethical and Distributive Considerations.”

7

Justice

In "The Lottery in Babylon," Jorge Luis Borges depicts a society that distributes all social benefits and burdens solely on the basis of a periodic lottery. Each person is assigned a social role such as slave, factory owner, priest, or executioner, purely by the lottery. This random selection system disregards criteria of distribution such as achievement, education, merit, experience, contribution, need, deprivation, and effort. The ethical and political oddity of the system described in Borges's story is jolting because assigning positions in this way fails to cohere with conventional conceptions and principles of justice. Borges's system appears capricious and unfair, because we expect valid moral principles to determine how social burdens, benefits, opportunities, and positions ought to be distributed.[1]

However, attempts to specify principles of justice for the many contexts of the distribution of health care and public health measures have often proved as inconclusive as the lottery method seems capricious. The construction of a unified theory of justice that captures our diverse conceptions and principles of justice in biomedical ethics continues to be controversial and difficult to pin down.

We start to work on these problems in this chapter by analyzing the terms *justice* and *distributive justice*. We then examine several general theories of justice pertinent to the distribution of health care. Later we examine problems of national and international health policy and consider enduring problems of social justice, including the nature of fair opportunity and unfair discrimination in health care, issues of vulnerability and exploitation of human subjects in research, the defensibility of the claim that we have both a right to health care and a right to health, several problems of global justice, the place of allocation and priority setting in health policy, and proper criteria of rationing health care in circumstances of scarcity.

The Concept of Justice and Principles of Justice

The terms *fairness*, *desert* (what is deserved), and *entitlement* have been used by philosophers as a basis on which to explicate the term *justice*. These accounts interpret justice as fair, equitable, and appropriate treatment in light of what is due or owed to affected individuals and groups. The term *distributive justice* refers to fair, equitable, and appropriate distribution of benefits and burdens determined by norms that structure the terms of social cooperation.[2] Its scope includes policies that allot benefits and burdens such as property, resources, taxation, privileges, opportunities, food distribution, jury service, and service as a research subject.

A compelling example of difficulties in determining the scope of distributive justice appears in the recent history of research involving human subjects. Until the 1990s, the paradigm problem in ethical assessment of research was the risks and burdens of research and the need to protect subjects from harm, abuse, and exploitation, especially when research offers no prospect of direct therapeutic benefit to the subjects and unfairly burdens a specific class of subjects. However, a paradigm shift occurred in the 1990s, in part because of the interest of patients with HIV/AIDS in gaining expanded access to new, experimental drugs both within and outside of clinical trials. The focus shifted to the possible benefits of clinical trials. As a result, justice as fair access to research—both participation in research and access to the results of research—became as important as protection from harm and exploitation.[3] (See further our discussion of programs and policies of expanded access and continued access to investigational [experimental] products such as drugs and medical devices in Chapter 6, pp. 224–27.)

No single moral principle is capable of addressing all problems of justice. Accordingly, in this chapter we discuss several principles of justice and consider how they can be balanced and specified in contexts of health care and public health. We argue that conditions of scarcity sometimes force a society to make tragic choices, and in the process, even valid principles of justice may be justifiably infringed, compromised, or sacrificed.[4]

We start first with a basic *formal* principle and then turn to principles that have been proposed as *material* principles, that is, substantive principles of justice.

The Formal Principle of Justice

Common to all theories of justice is a minimal requirement traditionally attributed to Aristotle: Equals must be treated equally, and unequals must be treated unequally. This principle of formal justice—sometimes called the principle of formal equality—is "formal" because it identifies no particular respects in which equals ought to be treated equally and provides no criteria for determining whether two or more individuals are in fact equals. It merely asserts that persons equal in whichever respects are relevant should be treated equally.

This formal principle lacks all substance. That equals ought to be treated equally provokes no debate, but significant problems surround judgments about what constitutes an equal and which differences are relevant in comparing individuals or groups. As a matter of human rights (see our account of rights in Chapter 9, pp. 400–9), all citizens in a political state should have equal political rights, equal access to public services, and equal treatment under the law, but how far do such principles of equality extend? Consider the following situation: Virtually all accounts of justice in health care hold that delivery programs and services designed to assist persons of a certain class, such as the poor, the elderly, pregnant women, children, and the disabled, should be made available to all members of that class. To deny benefits to some when others in the same class receive benefits is unjust, but is it also unjust to deny access to equally needy persons outside the delineated class, such as workers with no health insurance? How do we determine which classes, if any, should be designated? Answers require *material principles* of justice.

Material Principles of Justice and Morally Relevant Properties of Persons

Principles that specify the relevant characteristics for equal treatment are *material* because they identify the substantive properties for distribution. A relatively simple example is the principle of need, which dictates that essential social resources, including health care, should be distributed according to need. To say that a person needs something is to say that, without it, that person will suffer a harm, or at least be detrimentally affected. However, we are not required to distribute all goods and services to satisfy all needs, such as needs for athletic equipment and cell phones. Presumably our obligations are limited to fundamental needs for essential resources. To say that someone has a fundamental need is to say that the person will be harmed or detrimentally affected in a fundamental way if the need is not met. For example, the person might be harmed through malnutrition, bodily injury, or nondisclosure of critical information. (See our discussion of harm in Chapter 5, pp. 158–59.)

If we were to analyze the notion of fundamental needs further, we could progressively specify and shape the material principle of need into a public policy for purposes of distribution—for example, a public policy regarding who has a right of access to emergency rooms in hospitals, who may and who may not be on a list for an organ transplant, and the like. For the moment, however, we are emphasizing only the significance of accepting a principle of need as a valid material principle of justice. This principle is only one among several plausible material principles of justice. If, by contrast, one were to accept only a principle of free-market distribution as a valid principle of justice, then one would oppose a principle of need as a basis for public policy. All public and institutional policies based on distributive justice ultimately derive from the acceptance or rejection of some set of material principles and some procedures for specifying, refining, or balancing them; but determining precisely which principles are valid in which contexts continues to be a major problem for both theories of justice and public policy.

Material principles identify morally relevant properties that persons must possess to qualify for particular distributions, but theoretical and practical difficulties confront the justification of allegedly relevant properties. Tradition, convention, and moral and legal principles point to relevant properties in some cases, but it is often appropriate either to institute a new policy that establishes relevant properties where none previously existed or to revise entrenched criteria. For example, nation-states need to establish a policy about whether residents who are noncitizens will be allowed on waiting lists for transplants of deceased donor organs. The government must

decide whether citizenship is a relevant property and, if so, on which basis, in which ways, and with which exceptions.

Courts have sometimes mandated policies that revise entrenched notions about morally relevant properties. For example, the United States Supreme Court decided in the case of *Auto Workers v. Johnson Controls, Inc.*[5] that employers cannot legally adopt "fetal protection policies" that specifically exclude women of childbearing age from a hazardous workplace, because these policies unfairly discriminate based on the morally irrelevant property of gender. Under the policy that was challenged, fertile men could choose whether they wished to assume reproductive risks, whereas fertile women could not. The majority of justices held that this policy used the irrelevant property of gender despite the fact that mutagenic substances affect sperm as well as eggs.

Material Principles in Theories of Justice

Material principles are essential components of general theories of justice. We will introduce problems of distributive justice using this approach, turning first to what we call four traditional theories. We then consider two theories that attend closely to the value of health and health care in a theory of justice. Our main interest in examining these six theories is to call attention to various general principles that help us think through problems of justice in different contexts of biomedical ethics.

The four traditional theories are (1) *Utilitarian* theories, which emphasize a mixture of criteria for the purpose of increasing or maximizing human welfare and public utility; (2) *libertarian* theories, which emphasize individual rights to social and economic liberty, while invoking fair procedures as the basis of justice rather than substantive outcomes such as increases of welfare; (3) *communitarian* theories, which underscore principles of justice derived from conceptions of the good developed in moral communities; and (4) *egalitarian* theories, which emphasize equal access to the goods in life that every rational person values, often invoking material criteria of need and equality. The two theories specifically concerned with the value of health in the theory of justice are (5) *capability* theories, which identify capabilities such as the capability to be healthy that are essential for a flourishing life and identify ways social institutions can and should protect and promote capabilities; and (6) *well-being* theories, which emphasize core dimensions of human welfare and what is required nationally and globally to realize them.

Each of these six theories articulates a general, notably abstract, material principle of distributive justice:

1. 1. To each person according to rules and actions that maximize social utility (utilitarianism)
2. 2. To each person a maximum of liberty and property resulting from the exercise of liberty rights and participation in fair free-market exchanges (libertarianism)
3. 3. To each person according to principles of fair distribution derived from conceptions of the good developed in moral communities (communitarianism)
4. 4. To each person an equal measure of liberty and equal access to the goods in life that every rational person values (egalitarianism)
5. 5. To each person the means necessary for the exercise of capabilities essential for a flourishing life (capability theories)
6. 6. To each person the means necessary for the realization of core elements of well-being (well-being theories)

No obvious barrier prevents acceptance of more than one of these principles as valid—perhaps all six—in a pluralistic theory of justice. However, these principles are considered competitive in much of the literature on general theories of justice. To retain all six, one would have to argue that each of these material principles identifies a prima facie obligation whose weight cannot be assessed independently of particular goods and domains in which they are applicable, and then one would have to show how these principles can be rendered coherent in a pluralistic theory of justice.

Many and perhaps most societies invoke more than one of these material principles in framing public policies for different contexts. For example, the resources available for public health programs and for women's and

children's health programs are often distributed on the basis of either social utility or individual need to have a state of health protected or restored. Salaries and the higher incomes of some persons are sometimes allowed and even encouraged on grounds of free-market wage exchanges and competition. The resources needed for a basic education, for overcoming poverty, and for a decent level of health care are often distributed either equally to all citizens or as needed for citizens to achieve a basic level of well-being; and jobs and promotions in many sectors are awarded on the basis of demonstrated achievement and merit, as assessed by criteria in particular communities.

TRADITIONAL THEORIES OF JUSTICE

Theories of distributive justice link the morally relevant properties of persons to morally justifiable distributions of social benefits and burdens. By the last quarter of the twentieth century it became clear that the four traditional theories we will now examine had emerged as the most widely discussed theories of justice. We do not suggest that these theories are of equal importance, and we make no attempt to rank one over the others. In referring to them as "traditional," we are not signaling that they have a lower status, as if they were merely a matter of tradition and not currently defensible. Egalitarianism—the third theory—has been widely discussed and has probably become the most influential type of theory in philosophy over the last few decades. It remains the starting point for a great many writers on distributive justice. Egalitarianism is also the logical transition point from "traditional" theories to the more recent theories that we treat in the following section, because the commitment to proper distributions of health care and health programs in these recent theories show a significant egalitarian influence.

Utilitarian Theories

Utilitarian theories, which rose to prominence in the nineteenth century at the hands of John Stuart Mill and Jeremy Bentham, are treated as general moral theories in Chapter 9 (pp. 388–94). Principles of distributive justice, in particular, are presented in utilitarian theories as among several principles and rules that maximize utility or welfare. In this theory, any standard or rule of justice must be grounded in the principle of utility, which requires that we seek to produce the maximal balance of positive value over disvalue—or the least possible disvalue, if only undesirable results can be achieved.

Mill maintained that justice is the name for the paramount and most stringent forms of obligation fixed by the principle of utility.[6] However, the idea of maximizing utility is imprecise and has led to issues regarding which welfare functions should be maximized. In effect, all health benefits stand to improve welfare—for example, nutritious foods, clean water, hygiene, annual medical physical examinations, and public health measures. A utilitarian with a practical account of justice will explain how welfare is to be understood and how to weight conditions of welfare in the system.

Typically, utilitarian obligations of justice establish correlative rights for individuals that should be enforced by law (see our account of the correlativity of rights and obligations in Chapter 9, pp. 405–6). These rights are strictly contingent upon social arrangements that maximize net social utility. Human rights and principles of obligation have no other basis in this theory than utility maximization. Disputes abound among utilitarians as to whether rights have a meaningful place in utilitarian theory, but if a system of rights such as an international code of the rights of research subjects is justified entirely on the grounds that its existence will maximize social utility, utilitarians have no clear way to object to this defense of the rights.

However, as even many utilitarians point out, moral problems surround the use of utilitarian principles to justify rights such as the right to health care and the rights of human subjects. Rights grounded in justice could be viewed as having a tenuous foundation when they rest on overall utility maximization, because the balance of social utility could change at any time, and then the rights would also change. One coherent utilitarian outlook is that as conditions of social utility change, so the range of protected rights may change. For example, legal rights to health care in the United States have been limited to a few populations, notably the poor, the elderly, and military veterans; but conditions of social utility could shift so that a society has an obligation—flowing from

the principle of utility—to provide every citizen with a decent level of health care, which would be tantamount to conferring a right to health care for utilitarian reasons.

Although utilitarian theories face serious challenges as general theories of justice, the goal of maximizing social utility can be helpful in formulating just health policies in publicly supported institutions, especially when the policies are formulated using sensible cost-benefit or risk-benefit analysis, as we note later in this chapter and also in Chapter 6 (see pp. 243–52).

Libertarian Theories

Libertarian theories, as moral theories, date at least to early modern conceptions of natural rights, perhaps most notably to passages in John Locke's philosophy, which recognizes "just and natural rights" to liberty.[7] These theories are both general moral accounts and accounts of justice because they state the general duties that all members of society owe to one another, usually conceived as duties to respect liberty and to enforce individual liberty rights by coercive power when necessary. A libertarian interpretation of justice focuses not on public utility or on acting to meet the health and welfare needs of citizens, but on the unfettered operation of fair procedures and transactions under conditions of law and order.

Robert Nozick's work has for several decades been the most sophisticated and influential libertarian philosophical theory. He argues for a theory of justice in which government action is justified if and only if it protects citizens' liberty and property rights.[8] This theory of justice affirms individual liberty rights rather than a system that creates patterns of distribution in which governments collect taxes and redistribute the wealth originally acquired by persons in the free market. Governments act coercively and unjustly when they tax the wealthy at a progressively higher rate than the rate imposed on those who are less wealthy and then use the proceeds to underwrite state support of the indigent through welfare payments and unemployment compensation.

Nozick proposes three and only three principles of justice, all centered on private property rights: justice in acquisition, justice in transfer, and justice in rectification. No pattern of just distribution exists independent of free-market procedures of acquiring property, legitimately transferring that property, and providing rectification for those whose property was illegitimately taken or who otherwise were illegitimately obstructed in the free market. In this theory justice consists in the operation of just *procedures*, not in the production of just *outcomes* such as an equal distribution of health resources. The theory recognizes no fundamental welfare rights, and therefore no rights or justified claims to health or to health care can be based on justice. However, libertarians do not oppose utilitarian or egalitarian patterns of distribution if these patterns are freely chosen by all participants affected. Any distribution of goods, including public health measures and health care, is just and justified if and only if individuals in the relevant community freely choose it as public policy.

United States public policy has traditionally accepted a system that approximates a libertarian ideal according to which distributions of health insurance and health care are best left to a material principle of ability to pay for insurance and medical care, supplemented by voluntary charitable acts, institutions such as charitable hospitals, and employer-funded health insurance. Under this conception, a just society protects rights of property and liberty, allowing all persons the freedom to improve their circumstances and protect their health on their own initiative. Health care is not a right, the ideal system of health insurance is privatized, and charitable care institutions are nonprofit and untaxed.

Egalitarian Theories

Egalitarian theories have a history at least as old as religious traditions holding that all humans must be treated as equals because they are created as equals and have equal moral status. In moral and political philosophy, at least since Locke and other seventeenth-century writers, egalitarian thought has had a large presence. These theories explicate the idea of equality in terms of treating persons as equals *in certain respects*. No prominent egalitarian theory has contained a distributive principle that requires equal sharing of *all* social benefits and

burdens to all persons. The dominant egalitarian theories identify basic equalities for all persons while permitting some inequalities.

Rawls's celebrated egalitarian theory holds that "what justifies a conception of justice is not its being true to an order antecedent and given to us, but its congruence with our deeper understanding of ourselves and our aspirations."[9] In Rawls's account a theory of justice starts with considered judgments of equal respect for persons and fairness, which are specified in the theory to establish principles of justice. He argues that impartial persons would agree on two fundamental principles. The first principle requires that each person be permitted the maximum amount of basic liberty compatible with a similar measure of liberty for others. The second principle requires that social inequalities allowed in a theory must satisfy two conditions: (1) the first condition stipulates that inequalities in social primary goods (including, for example, inequalities in income, rights, and opportunities) are allowable, but only if allowing these inequalities benefits everyone ("the difference principle"); (2) the second condition requires that social offices and positions be open to all under circumstances of fair equality of opportunity (a fair opportunity principle, as we treat it later in this chapter on pp. 282–86).[10] Rawls considers nations and social institutions just (in liberal nation-states) if and only if they conform to each of these basic principles. He neither states how large the inequalities in income, rights, and opportunities might be nor speculates about how much better off the least advantaged must be under the difference principle. This position leaves uncertain how far the difference principle pushes in the direction of allowing inequalities, a tricky challenge for Rawlsians.

Although Rawls never pursued the implications of his theory for health policy in particular, others have done so. In an influential interpretation and extension, Norman Daniels argues for a just health care system based primarily on these principles, with a special emphasis on what Rawls called "fair equality of opportunity." Daniels sees health care needs as special and maintains that fair opportunity is central to any acceptable theory of justice. Social institutions affecting health care distribution should be arranged, insofar as possible, to allow each person to achieve a fair share of the normal range of opportunities present in that society.

Daniels's theory, like Rawls's, recognizes a positive societal obligation to reduce or eliminate barriers that prevent or reduce fair equality of opportunity, an obligation that extends to programs to correct or compensate for disadvantages. Daniels views disease and disability as undeserved restrictions on persons' opportunities to realize basic goals. Health care is needed to achieve, maintain, or restore adequate or "species-typical" levels of functioning so that individuals can realize basic goals. A health care system designed to meet these needs should be structured to prevent disease, illness, or injury from reducing the range of opportunity open to individuals; and the allocation of health care resources should be structured to ensure justice through fair equality of opportunity.[11]

This Rawls-inspired theory has far-reaching egalitarian implications for national health policies and perhaps for international health policies as well: Each member of society, irrespective of wealth or position, has equal access to an adequate, although not maximal, level of health care—the exact level of access being contingent on available social resources and public processes of decision making.

Communitarian Theories

So-called communitarian theories of justice can and have laid claim to traditions traceable to Aristotle and possibly to philosophers as different as Georg Wilhelm Friedrich Hegel and David Hume. However, only a small number of philosophers have self-identified as "communitarian," a label that collects a variety of theories focused on the relationship between individuals and their social embeddedness in communities, especially in the way communities shape individuals and construct their roles. What unites these theories under the somewhat artificial label of "communitarianism" (artificial since several leading "proponents" of this theory typically do not use, and some even reject, this label in their writings[12]) is the respect for and high value placed on the moral and political commitments found in communities and their traditions and practices. Individualism and individual rights that disrupt community goals have no place—or at least a radically reduced place—in these theories.

In recent forms, these theories critically react to "liberal theories of justice" such as those of Mill and Rawls, and, somewhat secondarily, to the libertarian theories of Nozick and others. Rawls's so-called political liberalism has been a special target of communitarian writers, who see societies constructed on these liberal foundations as lacking in a commitment to the general welfare, to common purposes, and to education in citizenship. Social conventions, traditions, loyalties, and the social nature of life and institutions figure prominently in many communitarian theories.[13] These theories particularly reject claims of the priority of the individual over the common good. Charles Taylor's challenge is straightforward: He argues that claims of the priority of individual rights over communal decision making are premised on a conception of the human good (e.g., the good of autonomous moral agency), as if individuals are isolated atoms existing independently of communities. Any theory of the place of autonomy that suggests a strong sense of independence, Taylor argues, is unacceptable if developed in the absence of the family and other community structures and interests.[14]

Communitarians see principles of justice as pluralistic, deriving from as many different conceptions of the good as there are diverse moral communities. What is owed to individuals and groups depends on community-derived standards.[15] As an example of communitarians' promotion of the common good in biomedical ethics, consider their difference from libertarians and others about policies for obtaining organs for transplantation from deceased persons. Based on principles of individual rights, all states in the United States adopted the Uniform Anatomical Gift Act in the late 1960s and early 1970s. This act gives individuals the right to donate their organs after death or to oppose donation, thereby blocking any possible familial decision to donate. However, if the decedent did not explicitly say "no" to donation, the next of kin then has the right to donate the decedent's organs.

Some communitarians challenge whether the individual's right to donate voluntarily is the primary consideration. A robust communitarian policy supports the *routine removal* of organs in the absence of registered objections by a donor. Arguments for this policy stress either the individual's obligation to donate to help others or society's ownership of organs of deceased individuals. Some communitarians argue for this policy on grounds that members of a society should be willing to provide others with objects of lifesaving value when they can do so at no cost to themselves. Other even more robust accounts recommend policies of routine removal that assume communal, rather than individual or familial, ownership of body parts of deceased persons.[16]

An emphasis on the community and the common good also appears in recommended policies for the allocation of health care. According to Daniel Callahan's avowedly communitarian account, we should enact public policy from a shared consensus about the good of society rather than on the basis of individual rights. Liberal assumptions about government neutrality should be reduced, and society should be free to implement a substantive concept of the good. For Callahan, the basic question is "What is most conducive to a good society?" The basic question is not, as he thinks many in bioethics assume, "Is it harmful, or does it violate autonomy?"[17]

TWO THEORIES CLOSELY CONNECTED TO THE VALUE OF HEALTH

Since roughly the end of the twentieth century, two innovative theories have reoriented discussions about justice in health policy and biomedical ethics. Both theories are inspired by Rawls and can be described as egalitarian, but they cannot be accurately described as fundamentally Rawlsian. They have also been deeply influenced by Aristotle's moral theory, especially his views about the role and importance of states of human flourishing.[18] This section examines these two theories.

Capability Theories

An approach known as capability theory presents an account of justice that starts from the premise that the opportunity for individuals to reach states of proper functioning and well-being is of basic moral significance and that the freedom to achieve these states should be analyzed in terms of the capabilities of individuals—that is, persons' powers or abilities to act and become what they value. In this theory, the quality of their lives is

contingent on what they are able to achieve, and lives well lived are those in which individuals sustain and exercise a group of identifiable core capabilities.[19]

This theory was pioneered by Amartya Sen as a way to address problems of welfare, justice, and human rights. For him, an individual's welfare is not determined by traditions that look to intrinsic goods or primary goods—as in utilitarian theory and in Rawls's theory of social primary goods, which holds that these goods are the ones that any rational person would want irrespective of anything else they might want.[20] Rather, Sen looks to actual opportunities for living well so that individuals can carry out or be what they value. "In contrast with the utility-based or resource-based lines of thinking," Sen writes, "individual advantage is judged in the capability approach by a person's capability to do things he or she has reason to value." In this account, the major way we assess the aggregate moral progress of human societies is by the expansion of human capabilities and the reduction of inequalities through the development of capabilities in those societies.[21]

This capability approach has been developed with an Aristotelian touch—and often with direct relevance to biomedical ethics—by Martha Nussbaum.[22] She uses the theory to address "social justice" and the "frontiers of justice"—the latter including justice for the disabled, the global poor, and nonhuman animals. Nussbaum's theory holds that a minimal level of social justice requires making available "to all citizens … the following ten core 'capabilities,'" which she calls "the central human capabilities":[23]

1. 1. *Life.* Being able to live a normal life without dying prematurely or existing in a reduced state making life not worth living
2. 2. *Bodily health.* Being able to have good health, nutrition, and shelter
3. 3. *Bodily integrity.* Being able to move freely, to be secure against violence, and to have opportunities for sexual satisfaction and reproductive choice
4. 4. *Senses, imagination, and thought.* Being able to use these capacities in an informed and human way aided by an adequate and diverse education and in a context of freedom of expression
5. 5. *Emotions.* Being able to have emotional attachments to persons and things so that one can love, grieve, and feel gratitude without having one's emotional development blunted by fear, anxiety, and the like
6. 6. *Practical reason.* Being able to form a conception of the good and to critically reflect in planning one's life
7. 7. *Affiliation.* Being able to live meaningfully in the company of others, with self-respect and without undue humiliation
8. 8. *Other species.* Being able to live with concern for animals, plants, and nature generally
9. 9. *Play.* Being able to play and enjoy recreational activities
10. 10. *Control over one's environment.* Being able to participate as an active citizen in political choices pertaining to one's life and property

Each capability is essential for a human life not to be impoverished below the level of the dignity of a person, and each capability forms the basis of a human right or entitlement. Our natural *basic* capabilities should be developed, in this theory, in a manner that generates *trained* capabilities. For example, we innately have capacities for speech, learning, and free action, which can then be developed into more advanced capabilities such as literacy, job skills, and knowledge about how to avoid poverty and disease.

In Nussbaum's account, these capabilities are essential to flourishing and must, as a matter of justice, be socially sustained and protected. She claims that "*all ten of these plural and diverse ends are minimum requirements of justice*, at least up to [a] threshold level."[24] Justice requires that we, as a society, see to it that all ten capabilities are secured for all citizens to the designated threshold level; and we must ensure that actors and conditions in the world do not interfere with individuals' development of their core capabilities or block political participation in a way that stunts or harms them.

Society also sometimes must equip persons with capabilities that include provision of the resources necessary for living appropriately, such as food, education, nondiscriminatory institutions, and health care. This approach focuses on putting persons in circumstances in which they are enabled to set their own goals and live as they

choose, thereby bringing the theory close to theories of individual, social, and political freedom. Nussbaum also brings an egalitarian dimension to her theory by insisting that "the political entitlements of all citizens are equal and the same."[25]

In addressing what Nussbaum calls the "frontiers of justice," her theory is remarkably broad, covering not only human capabilities and functioning for disabled and socially oppressed persons but also for nonhuman animals. Treating an individual—human or nonhuman—justly requires, negatively, *not obstructing* the individual's attempts at flourishing through acts of coercion, violence, or cruelty and also requires, positively, support of efforts to flourish.[26] The inclusion of nonhumans renders this theory of justice extremely demanding, perhaps as bold and ambitious as any theory of justice ever devised. It is also what we will later discuss as a global theory that extends "justice to all world citizens, showing theoretically how we might realize a world that is just as a whole" by providing the "necessary conditions for a decently just society."[27]

Well-Being Theories

Capability theories are centered on the abilities, opportunities, and forms of freedom requisite for well-being, but a recent general theory that is closely aligned with biomedical ethics focuses on *well-being itself*, rather than on *capabilities* for well-being. In this account the focus is not on individuals being able to pursue states of affairs if they so choose. The focus is on ensuring that everyone experiences well-being in ways commensurate with a decent life. Utilitarianism could be read as this type of well-being or welfare theory, but in this subsection we concentrate on the rare case of a theory—a theory originally devised by Madison Powers and Ruth Faden—that is explicitly directed at public health and health policy.

They start with a basic premise: "Social justice is concerned with human well-being." It is concerned not merely with capabilities for well-being or with a single form of well-being such as health. They argue that a theory of social justice should be concerned centrally with six core elements of well-being:[28]

1. 1. Health
2. 2. Personal Security
3. 3. Knowledge and Understanding
4. 4. Equal Respect
5. 5. Personal Attachments
6. 6. Self-determination

This list of core elements, or dimensions, of well-being may seem similar to Nussbaum's list of capabilities—for example, "personal attachment" resembles Nussbaum's "affiliation"—but Powers and Faden emphatically reject the language of *capabilities* as the best way to express the basic notions in a theory of justice.[29] They regard being healthy, being secure, and being respected as desirable *states of being*. We want not merely the *capability* to be secure and healthy, but to *be* secure and healthy. Justice in this theory is concerned with society's achievement of well-being for its members, not solely with *securing basic capabilities* that enable us to pursue states of well-being. In the case of the core element of health, the major goal is instantiation of both the right to *health* and the right to *health care*.

The "job of justice"[30] is to secure the core elements of well-being in all of the six dimensions for each person in every society, including the global society. Each of the six dimensions is an independent concern of justice, but they also interact with each other. The justice of health policies in both particular societies and the global order can be judged not only by how well they serve to secure health, but also by their effects on the other core elements of well-being.

Powers and Faden regard the goal of justice as that of securing well-being and human rights and also ensuring against unfair relations of power and advantage. Theirs is a structural theory of justice—a notion they elaborate in detail in their second book, *Structural Injustice: Power, Advantage and Human Rights*. Their focus on the relationship between well-being and human-rights norms and fairness norms highlights the role that poverty and

unfair disparities in power and advantage play in causing and perpetuating poor health and injustice in countries around the globe.

Powers and Faden see the goal of egalitarian justice as the reduction of inequality in the world as we encounter it—a world characterized by profound inequalities in well-being, resources, power, and advantage. Although health is only the first of the six core elements of well-being, Powers and Faden argue that the moral justification for health policies depends as much on the other five dimensions of well-being as it does on health—a critical feature of their theory. They argue that the absence of any of the other five core elements can be destructive to health. A constellation of inequalities can systemically magnify and reinforce initial conditions of ill health, creating ripple effects that impact other dimensions of well-being. The interactive effects include poor education and lack of respect, which can affect core forms of reasoning and health status. Social structures can compound these adverse effects. The result is a mixture of interactive and cascading effects that requires urgent attention from the point of view of justice.[31] The job of justice is to correct these defects by making the six core elements of well-being embedded values in social policy.

Conclusion

The six theories of justice we have considered can be expected to succeed only partially in bringing coherence and comprehensiveness to our multilayered and sometimes fragmented conceptions of social justice. Policies governing health care access and distribution in many nations provide excellent cases of the problems that confront these theories. Many countries seek to make available high-quality health care for all citizens while protecting public resources through cost-containment programs and respecting the choices of patients and clinicians. Their policies promote the ideal of equal access to health care and good health for everyone, including the indigent, while maintaining aspects of a competitive, free-market environment. These goals of quality health care, equal access, free choice, social efficiency, and good health are all laudable (we here presume without argument), but they are also difficult to render coherent in a social system and in a theory of justice. Pursuing one goal may function to undercut another.

It is likely that there has never been a political state or a world order fashioned *entirely* on one and only one of the six theories of justice that we have now discussed. Some commentators see these theories as having the weakness of Plato's ideal state in the *Republic*: They provide models but not truly practical instruments. This skeptical caution is prudent, but it can lead to an underevaluation of the moral implications and force of the six theories of justice we have considered. Intelligent use of the principles of justice in these theories has practical significance for biomedical ethics as well as for health and health care in political states and the global community, as we show in the remainder of this chapter. We will not attempt to assess the relative merits of these theories. Rather, we use them as resources, with special attention to recent egalitarian thinking and proposals about the distribution of health care and public health resources.

FAIR OPPORTUNITY AND UNFAIR DISCRIMINATION

The rule of fair opportunity is among the most influential features of egalitarian thinking in bioethics. We begin this section with the question, "What kind of fair opportunity does justice require?" We consider first the properties that have often served, *unjustly*, as bases of social distribution. These properties include gender, race, IQ, linguistic accent, ethnicity, national origin, and social status. In anomalous contexts, such as casting for film or theater, these typically irrelevant properties may become relevant and acceptable, though still challengeable even in these domains. However, general rules such as "To each according to gender" and "To each according to IQ" are unacceptable as prima facie material principles of justice. These properties are irrelevant and based on differences for which the affected individual is not responsible. Basing actions or policies on them is emphatically discriminatory.

The Fair-Opportunity Rule

The fair-opportunity rule descends from Rawls's conditions of fair equality of opportunity. The rule asserts that individuals should not receive social benefits on the basis of undeserved advantageous properties and should not be denied social benefits on the basis of undeserved disadvantageous properties, because these individuals are not responsible for these properties. Properties distributed by the lotteries of social, psychological, and biological life do not provide grounds for morally acceptable discrimination between persons in social allocations if people do not have a fair chance to acquire advantaging properties or overcome disadvantaging properties.

The goal of supplying all citizens with a basic *education* raises moral problems analogous to these problems of *justice in health care*. Imagine a community that offers a high-quality education to all students with basic abilities, regardless of gender or race, but does not offer a comparable educational opportunity to students with reading difficulties or mental deficiencies. This system is unjust. The students with disabilities lack basic skills and require special training to overcome their problems to the extent they can. They should receive an education suitable to their needs and opportunities, even if it costs more. The fair-opportunity rule requires that they receive benefits that will ameliorate the unfortunate effects of life's lottery. By analogy, persons with functional disabilities lack critical capacities and need health care to reach a suitable level of function and have a fair opportunity in life. When persons are not responsible for their disabilities, the fair-opportunity rule demands that they receive help to reduce or overcome the unfortunate effects of life's lottery of health.

Fair Opportunity as a Rule of Redress: Mitigating the Negative Effects of Life's Lotteries

Numerous properties might be disadvantageous and undeserved—for example, a squeaky voice, an ugly face, inarticulate or halting speech, an inadequate early education, malnutrition, and disease. But which undeserved properties create a right *in justice* to some form of assistance in ameliorating their disadvantaging conditions?

A strong claim is that virtually all abilities and disabilities are functions of what Rawls calls the natural lottery and the social lottery. "Natural lottery" refers to the distribution of advantageous and disadvantageous genetic properties, and "social lottery" refers to the distribution of assets or deficits through family property, school systems, tribal affiliation, government agencies, and the like. It is conceivable that all talents, disabilities, and disadvantaging properties result from sources such as heredity, natural environment, family upbringing, education, and inheritance. Even the ability to work long hours, the ability to compete, and a warm smile may be biologically, environmentally, and socially engendered. If so, talents, abilities, and successes are not to our credit, just as genetic disease is acquired through no fault of the afflicted person.

Rawls uses fair opportunity as a *rule of redress*, as do we. To overcome undeserved disadvantaging conditions, whether they derive from the natural lottery or the social lottery, the rule of fair opportunity demands compensation for disadvantages. The full set of implications of this theory have never been made entirely clear, but Rawls's conclusions are demanding:

> [A free-market arrangement] permits the distribution of wealth and income to be determined by the natural distribution of abilities and talents. Within the limits allowed by the background arrangements, distributive shares are decided by the outcome of the natural lottery; and this outcome is arbitrary from a moral perspective. There is no more reason to permit the distribution of income and wealth to be settled by the distribution of natural assets than by historical and social fortune. Furthermore, the principle of fair opportunity can be only imperfectly carried out, at least as long as the institution of the family exists. The extent to which natural capacities develop and reach fruition is affected by all kinds of social conditions and class attitudes. Even the willingness to make an effort, to try, and so to be deserving in the ordinary sense is itself dependent upon happy family and social circumstances.[32]

Current social systems of distributing benefits and burdens would undergo massive revision if this approach were accepted and applied in social policy. Instead of permitting broad inequalities in access to health care and quality of care—based on employer contributions, wealth, celebrity status, and the like—justice would be achieved only if opportunity-reducing inequalities are first addressed. At some point processes of reducing the

inequalities created by life's lotteries must end because of social limits of resources.[33] From this perspective, a strict fair-opportunity rule will be overly demanding unless it is carefully qualified.

Redressing Racial, Ethnic, Gender, and Social Status Disparities in Health Care

Many disparities in health care and research are based on race, ethnicity, gender, and social status and thus undermine fair opportunity. Health care goods and research risks have often been covertly distributed on the basis of these properties, resulting in a differential impact in many countries on the health of racial and ethnic minorities, women, and the poor.[34] Various studies in the United States indicate that African Americans, other minorities, women, and the economically disadvantaged have poorer access to various forms of health care and to valued research in comparison to white males. For example, gender and racial inequities in employment situations have an impact on job-based health insurance, and serving as research subjects often falls disproportionately on socially and economically disadvantaged individuals, who have low incomes or are homeless. Similarly, some forms of health care and research disproportionately benefit patients who are already socially and economically advantaged.

Numerous efforts have emerged to identify and address racial, ethnic, gender, and social-status disparities.[35] One controversy centers on disparities in rates of coronary artery bypass grafting (CABG) between white and black Medicare patients, as well as between male and female Medicare patients. Differences in use, which have been evident since the 1980s, cannot be entirely accounted for by differential need, and it remains unclear to what extent the rates can be explained by conditions such as physician supply, poverty, awareness of health care opportunities, reluctance among blacks and women to undergo surgery, and racial prejudice. One study found that, after controlling for age, payer, and appropriateness and necessity for CABG, African American patients in New York State still had significant access problems unrelated to patient refusals.[36] Moreover, racial disparities persist in outcomes of CABG, probably reflecting complex factors that may be difficult to identify and correct.[37]

Disparities have persisted in other areas as well, including the management of acute myocardial infarction and acute coronary syndromes,[38] in cholesterol control among patients with cardiovascular diseases, in cancer screening,[39] in the diagnosis and treatment of conditions such as colorectal cancer and glucose control for patients with diabetes,[40] and in pain care.[41] These disparities are among the unacceptable racial and ethnic differences across a range of medical conditions and health care services that lead to worse health outcomes and that were identified in the Institute of Medicine (now National Academy of Medicine) report on unequal access *Unequal Treatment: Confronting Racial and Ethnic Disparities in Health Care*.[42] Disparities in use do not always amount to injustices, but they require close scrutiny to determine their causes and to guard against injustice.[43]

Several complex factors are associated with disparities in health care, not to mention disparities in health, which raise even more challenging questions about causes and corrective measures. Some disparities in health care may result from implicit biases, even though the causal links may not be as direct as sometimes supposed. Consider disparities in pain care, for example. The history of the differential treatment of pain in the United States displays a variety of social and cultural beliefs and values, including some that lead to systematic undertreatment of African Americans for pain.[44] One study found that false beliefs about biological differences between whites and blacks (e.g., regarding the thickness of blacks' skin) correlated with assessments of the level of pain experienced by whites and blacks. These beliefs were not held only by laypeople but also by half of a sample of medical students and residents at a major medical center. Medical participants with these beliefs rated black patients' pain lower than white patients' pain and proposed less correct or less appropriate treatments.[45] This study and others indicate the importance of attention to both explicit and implicit bias as possible causal factors in disparities in health care that must be addressed along with culture, social structures, and policies. For instance, one vignette study found that implicit racial bias (as measured by Implicit Association Tests) predicted physicians' differential assessments and recommendations of white and black patients for thrombolysis to treat myocardial infarction.[46]

Renal transplantation provides an illuminating example of racial disparities and vigorous—and somewhat successful—efforts to achieve fairer policies and practices. In United States policy financial barriers play a less significant role in kidney transplantation than in most areas of health care. The reason is that the federal End-Stage Renal Disease (ESRD) Program ensures coverage for kidney dialysis and transplantation for virtually every citizen who needs them if their private insurance does not provide the coverage. However, concerns about costs are still a factor because immunosuppressant medications needed to maintain a transplanted kidney throughout the recipient's life are covered under the ESRD program for only three years. Evidence suggests that discrimination against blacks, other minorities, women, and the poor occurs leading up to and including referral to transplantation centers, where criteria for admission to waiting lists may vary considerably. For instance, black Americans are much less likely than white Americans to be referred for evaluation at transplant centers and to be placed on a waiting list or to receive a transplant.[47] Factors include delayed or limited access to health care, inadequate care and guidance through the system by health care professionals, and minorities' distrust of the system.

Once patients are admitted to the waiting list, the criteria for selecting recipients of deceased donor organs are public and are represented through point systems. Disputes continue regarding how much weight to give to different factors in the distribution of kidneys for transplantation, with particular attention to human lymphocyte antigen (HLA) matching. The degree of HLA match between a donor and a recipient affects the long-term survival of the transplanted kidney. However, assigning priority to tissue matching—and giving less weight to time on the waiting list and other factors—has been shown to produce disparate effects for minorities. Most organ donors are white; certain HLA phenotypes are different in white, black, and Hispanic populations; and the identification of HLA phenotypes is less complete for blacks and Hispanics. Yet nonwhites have a higher rate of end-stage renal disease and are also disproportionately represented on dialysis rolls. Blacks on the waiting list, on average, wait longer than whites to receive a first kidney transplant, if they receive one at all.

After extensive discussion and deliberation, including professional and public input, the United Network for Organ Sharing in 2003 changed its kidney allocation criteria to eliminate the priority given to HLA-B matching, a subset of HLA phenotypes, with the goal of reducing the disparity in deceased donor kidney transplants between African Americans and whites. The revised policy was defended on the grounds that it would resolve "the tension inherent in the current allocation policy by improving equity without sacrificing utility."[48] It succeeded in reducing the disparity: Before the policy change, African Americans had 37% lower rates of deceased donor kidney transplants, but after the change they had 23% lower rates.[49] Although the disparity was significantly reduced, it was not fully eliminated, perhaps because of unaddressed or unknown factors.

Further studies of renal transplantation have found reduced racial disparity in the outcomes of kidney transplantation in the United States, particularly in regard to three-year and five-year graft loss among African Americans following a transplant.[50] Others caution that racial disparities persist in kidney transplantation, emphasizing that in addition to transplant outcomes—the focus of many studies—major disparities remain in access to care, particularly as seen in delays in referral for transplant evaluation, in access to the national waiting list, and in receiving a kidney transplant. Another recent policy initiative that may help is a rule that allows time on the waitlist to begin when the patient's kidney function declines to a certain level even before the patient has started dialysis or been referred for a transplant.[51]

It is not clear whether, or to what extent, the policy change in HLA-B matching decreased the number of years of transplant function. Normatively, however, the common tension between maximizing utility and providing fair opportunity persists in this area as in many others. Critics have charged that organ allocation policies should seek to maximize the number of quality-adjusted life-years per transplanted organ instead of using disparate impact tests of policies to try to increase the access of racial or ethnic groups to transplantation.[52]

We conclude that developing an equitable policy of organ distribution and allocation is challenging and requires attending to and monitoring disparities, considering the disparate impact of different policies as well as their overall medical utility (producing the most overall patient welfare), and finding the best available ways to balance medical utility and fair opportunity.

VULNERABILITY, EXPLOITATION, AND DISCRIMINATION IN RESEARCH

We now turn to a different, yet related, set of moral and social problems about fair opportunity that derive from the vulnerability of human research subjects at risk of exploitation. We concentrate on the recruitment and enrollment in clinical research—primarily in pharmaceutical trials—of the economically disadvantaged.

By "economically disadvantaged," we mean persons who are impoverished, may lack significant access to health care, may be homeless, or may be malnourished, and yet possess the mental capacity to volunteer to participate in, for example, safety and toxicity (phase 1) drug studies. We consider in this section only persons who possess a basic competence to reason, deliberate, decide, and consent. Somewhere between 50% and 100% of research subjects who are healthy volunteers self-report that financial need or financial reward is their primary motive for volunteering.[53] We know little about the full extent of their involvement in research, just as we do not know the scope of the use of poor persons as research subjects.[54]

Vulnerability and Vulnerable Groups

The relevant literature has sometimes viewed the class of the economically disadvantaged and vulnerable as narrow, at other times as broad. The persons so classified may or may not include individuals living on the streets, low-income persons who are the sole financial support of a large family, persons desperately lacking access to health care, persons whose income falls below a certain threshold level, and so forth.

The notion of a "vulnerable group" was a major category in bioethics and health policy between the 1970s and the 1990s.[55] In the years thereafter it suffered from overexpansion because so many groups were declared vulnerable—from the infirm elderly, to the undereducated, to those with inadequate resources, to whole countries whose members lack rights or are subject to exploitation.[56] The language of "vulnerable groups" suggests that all members of a vulnerable group—for example, all prisoners and all poor people—are by category vulnerable, but for many groups a label covering all members of the group serves to overprotect, stereotype, and even disqualify members capable of making their own decisions.[57] "Vulnerable" is an inappropriate label for any class of persons when some members of the class are not vulnerable in the relevant respects. For example, pregnant women as a class are not vulnerable, although some pregnant women are. Accordingly, we will not here speak of the economically disadvantaged as a vulnerable group. Instead, we will concentrate on *vulnerabilities*.[58]

A tempting strategy to protect the interests of economically disadvantaged persons is to exclude them categorically from participation in research even if they are not categorically vulnerable. This remedy would eliminate the problem of unjust exploitation but also would deprive these individuals of the freedom to choose and would often be harmful to their financial interests. Nothing about economically disadvantaged persons justifies their exclusion, as a group, from participation in research, just as it does not follow from their status as disadvantaged that they should be excluded from any legal activity they wish to pursue. Despite the increased risks of exploitation experienced by the economically distressed, to exclude them categorically would be an unjust and paternalistic form of discrimination that might serve to further marginalize, deprive, or stigmatize them.

Undue Inducement, Undue Profit, and Exploitation

Other moral problems about enrolling the economically disadvantaged in research include undue inducement, undue profit, and exploitation. Some persons report feeling heavily pressured to enroll in clinical trials, even though their enrollment is classified as voluntary.[59] These individuals may be in desperate need of money. Attractive offers of money and other goods can leave a person with a sense of being constrained and having no meaningful choice but to accept research participation.

Constraining situations. These *constraining situations* are sometimes misleadingly termed *coercive situations*.[60] Here a person feels controlled by the constraints of a situation, such as severe illness or lack of food or shelter, rather than by the design or threat of another person. No coercion has occurred because no one has intentionally issued a threat to gain compliance or forced a person to consent. Still, persons feel "threatened" by their situations and sometimes feel compelled to prevent or ameliorate perceived harms of illness, powerlessness, or lack of resources. The prospect of another night of homelessness on the streets or another day without food can compel a person to accept an offer of research participation, just as such conditions can compel a person to accept an unpleasant or risky job that the person would otherwise not accept.

Undue inducement. In constraining situations, monetary payments and related offers such as shelter or food generate problems of justice commonly referred to as *undue inducement*, on the one hand, and *undue profit*, on the other. The "Common Rule" in the United States requires investigators to "minimize the possibility of coercion or undue influence," but it does not define, analyze, or explain these notions.[61] The bioethics and public policy literatures also do not adequately handle the problems.

Monetary payments seem unproblematic if the payments are offers that persons find welcome and do not want to refuse and the risks are at the level of everyday activities.[62] But inducements become increasingly problematic as (1) risks are increased, (2) highly attractive inducements are offered, and (3) the subjects' economic disadvantage is elevated. The problem of exploitation centers on whether solicited persons are situationally disadvantaged and lack viable alternatives, feel forced or compelled to accept offers that they otherwise would not accept, and assume increased risk in their lives. As these conditions are mitigated, problems of exploitation diminish and may vanish altogether. As these conditions are increased, problems of exploitation loom larger.

The presence of an irresistibly attractive offer is a necessary condition of "undue inducement," but this condition is not by itself sufficient to make an inducement undue. A situation of undue inducement must also involve a person's assumption of a serious risk of harm that he or she would not ordinarily assume. Pinpointing a precise threshold level of risk is difficult, but the level would have to be above the level of common job risks such as those of unskilled construction work. Inducements are not undue unless they are both above the level of standard risk (hence excessive in risk) and irresistibly attractive (hence excessive in payment) for persons in a constraining situation.

Undue profit. Undue inducements should be distinguished from *undue profits*, which occur from a distributive injustice of too small a payment to subjects by contrast to an irresistibly attractive, large payment. In the undue-profit situation, the subjects in research receive an unfairly low payment, while the sponsor of research garners more than is justified. Often, this state of affairs seems to be what critics of pharmaceutical research claim to be the case: Researchers approach potential subjects who are in a weak to nonexistent bargaining situation, constrained by their poverty, and offered an unjustly small amount of money and an unjustly low share of the benefits, while companies reap unseemly profits. If this description captures the moral concern, the basic moral problem is how to determine a nonexploitative, fair payment for service as a research subject, which might include benefits of successful research such as free provision of an experimental drug after the end of a clinical trial.

How should we handle these two moral problems of exploitation—undue inducement (unduly large and irresistible payments) and undue profit (unduly small and unfair payments)? One approach is to prohibit research that involves excessive risk, even if a good oversight system is in place. This answer is appealing, but we would still need to determine in each case what constitutes excessive risk, irresistibly attractive payment, unjust underpayment, and constraining situations—difficult and unresolved problems.

These problems resist a tidy solution. To avoid undue inducement, payment schedules must be kept reasonably low, perhaps approximating an unskilled labor wage. Even at this low level, payment might still be sufficiently large to constitute an undue inducement for some research subjects. As payments are lowered to avoid undue inducement, research subjects in some circumstances will be recruited largely or entirely from the ranks of the economically disadvantaged. Somewhere on this continuum the amount of money paid will be so little that it is exploitative by virtue of undue profits yielded by taking advantage of a person's misfortune. If payment scales

were increased to avoid undue profit, they would at some point become high enough to attract persons from the middle class. At or around this point, the offers would be excessively attractive, undue inducements for impoverished persons interested in the payments.[63] Addressing this dilemma generates deep social injustice if the pool of research subjects is composed more or less exclusively of the economically disadvantaged.

Finally, an important reason for caution about prohibiting research or about encouraging pharmaceutical companies to pull out of communities with a significant percentage of poor people is that payments for studies may be a vital source of needed funds for the economically disadvantaged and a way to build an infrastructure and create jobs in communities. Among the few readily available sources of money for some economically distressed persons are jobs such as day labor that may expose them to more risk and produce less money than the payments generated by participation in phase 1 clinical trials.[64] To deny these persons the right to participate in clinical research on grounds of potential exploitation can be paternalistic and demeaning, as well as economically detrimental. In many circumstances such treatment would be unjust, but in other circumstances it would not involve unjust practices.

National Health Policy and the Right to Health Care

Problems of justice in access to health care differ substantially in many parts of the world, but questions about who shall receive which share of a society's resources are at the center of the discussion almost everywhere. In this and later sections we examine controversies about appropriate national health policies, inequalities in distributions of care, rationing of health-related goods and services, and problems of global justice.

The primary economic barrier to health care access in many countries is the lack of adequate insurance or funding for care. In the United States in 2016, the percentage of people *without* health insurance coverage for the entire calendar year was 8.8%, or 28.1 million; the percentage of people with some form of health insurance coverage for *all* or *part* of 2016 was 91.2%.[65] Inadequate insurance can profoundly affect persons who are uninsured, uninsurable, underinsured, or only occasionally insured. Some problems of unfairness arise in the United States because of the system's reliance on employers for financing a huge portion of health insurance. Persons employed by medium- to large-sized employers are generally covered better and may be subsidized by tax breaks. When employed persons who are not covered become ill, taxpayers (and not free-riding employers) may pick up the cost. The financing of health care is also regressive. Low-income families pay premiums comparable to and often higher than the premiums paid by high-income families, and many individuals who do not qualify for group coverage pay dramatically more for the same coverage than those who qualify in a group. A dark cloud of injustice hangs over these circumstances.

A social consensus appears to exist in the United States that all citizens should be able to secure equitable access to health care, including insurance coverage. However, this consensus is content-thin regarding the role of government, methods of financing insurance and health care, how much insurance is sufficient, and the meaning of "equitable access." It remains unclear whether such a fragile consensus can generate a secondary consensus about how to implement a system of equitable access. Similar issues appear in many nations.

Arguments Supporting Rights to Health Care

Whether there is a right to health care has long been debated in public discourse around the world. Some countries maintain the more or less libertarian view that citizens have a right to purchase health insurance, whereas no government obligation exists to supply health care. By contrast, some countries with roughly utilitarian, egalitarian, or communitarian commitments hold that citizens have a right to health care that is government-guaranteed. Many countries adopt elements of rights to health care based on pieces of these several theories of justice. In this section, we explore which position is the most morally defensible.

Two influential arguments support a moral right to government-funded health care (but other good arguments may also be available): (1) an argument from collective social protection and (2) an argument from fair opportunity. Even if these are good arguments, as we think they are, how much in the way of health care and

public health should be funded by the government? This question about sufficiency of government support may be the most important practical question in all of biomedical ethics.

The first argument focuses on the similarities between individuals' health needs and other needs that governments have traditionally met. Threats to health are often relevantly similar to threats presented by crime, fire, and pollution. Collective actions and resources have conventionally been used to resist such threats, and many collective schemes to protect health exist in virtually all societies, including extensive programs of public health and environmental protection. Consistency suggests that critical health care assistance in response to threats to health should likewise be a collective responsibility. This argument by analogy appeals to coherence: If the government has an obligation to provide one type of essential service, then it must have an obligation to provide another relevantly similar essential service.

This argument has been criticized on grounds that government responsibilities are neither obligatory nor essential. However, this perspective is favored by few beyond those committed to a strong libertarianism. On each of the nonlibertarian theories of justice previously explicated, the argument from other comparable government services generates a public obligation to provide some level of goods and services to protect health. Nevertheless, relevant dissimilarities exist between the good of health care for individuals and programs that protect the general public through social goods such as public health measures. The argument from collective social protection therefore might seem to fail or at least to be incomplete.

However, additional premises supporting the right to health care are found in society's right to expect a decent return on the investment it has made in physicians' education, funding for biomedical research, and funding for various parts of the medical system that pertain to health care. This supplementary argument appeals to reciprocity: Society should give a proportional return on benefits received from individuals, with all alike sharing the burdens of taxation necessary to produce these benefits. A return to be expected on individuals' investments in a well funded health system based on taxes is protection of their personal health. The scope of protection extends beyond public health measures to access to physicians and the products of research.

This argument based on reciprocity has merit, but we cannot reasonably expect a direct individual return on all of our collective investments in a health system. Some investments are only for the purpose of discovering treatments, not for the provision or distribution of treatments once discovered. Even if a government funds drug research and regulates the drug industry, these activities may not justify the expectation that the government will subsidize or reimburse individuals' drug purchases. Accordingly, this first argument in support of a moral right to health care may secure only a right to a decent return to individuals on their contribution to society's investment, not a full return.

A second argument buttresses this first argument by appeal to the previously discussed fair-opportunity rule, according to which the justice of social institutions should be assessed by their tendency to counteract lack of opportunity caused by unpredictable misfortune over which the person has no meaningful control. The need for health care is greater among the seriously diseased and injured, because the costs of health care for them can be uncontrollable and overwhelming, increasingly so as their health status worsens. Insofar as injuries, diseases, or disabilities create profound disadvantages and reduce agents' capacity to function properly, justice as the provision of a fair opportunity requires that we use societal health care resources to counter these effects and give persons a fair chance to develop, maintain, restore, and use their capacities.[66]

The Right to a Decent Minimum of Health Care

One problem about the right to health-related goods and services is how to specify the entitlements the right requires us to recognize. One egalitarian approach proposes a right of *equal access* to health resources. At a minimum, this goal entails that all persons have a right not to be prevented from obtaining health care, but this thin right of access does not entail that others must provide anything in the way of goods, services, or resources. Some libertarians favor not providing anything from public funds, but their proposal is not supported by the other general theories of justice we have examined. A meaningful right of access to health care includes the right to obtain goods and services to which every entitled person has an equal claim. A demanding interpretation of

this right appears in the global-justice claim that everyone everywhere should have equal access to all goods and services available to anyone. Unless the world's economic systems are radically revised and economic resources are abundantly available, this conception of a human right that applies globally will remain a utopian ideal that is unlikely to be fully realized. Rights to health-related resources will likely always have severe limits (see our discussions below of "Setting Priorities" and "Rationing," pp. 300–13). But this conclusion does not diminish the importance of practical pursuits of a human right in various areas of public health and health care.

The right to a decent minimum of health care presents a more attractive goal—and, realistically, probably the only goal that can be achieved in most societies seeking to implement a right to health care.[67] This egalitarian goal is one of universal accessibility in a political community to fundamental health care and basic health-related resources. The standard conception is a two-tiered system of health care: enforced social coverage for basic and catastrophic health needs (tier 1), together with voluntary private coverage for other health needs and desires (tier 2). At the second tier, better services, such as luxury hospital rooms and optional, cosmetic dental work, are available for purchase at personal expense through private health insurance or direct payment. The first tier meets needs by universal access to basic services. This tier presumably covers at least public health protections and preventive care, primary care, acute care, and special social services for those with disabilities. This model of a safety net that protects everyone acknowledges that society's obligations are demanding, but not limitless.

The decent minimum offers a possible compromise among the theories of justice discussed previously, because it incorporates some moral premises that most theories stress or at least find acceptable. It guarantees basic health care for all on a premise of equal access while allowing unequal additional purchases by individual initiative, thereby mixing private and public forms of distribution. An egalitarian should see an opportunity to use an equal access principle and to embed fair opportunity in the distributional system. Utilitarians should find the decent-minimum proposal attractive because it serves to minimize public dissatisfaction, to promote social utility, and to permit allocation decisions based on cost-effectiveness analysis. Similarly, supporters of a capability theory or a well-being theory can appreciate the likelihood of increases in the capability of many to afford better quality care and achieve higher levels of health. The libertarian may dislike these outcome-oriented approaches but still should see a substantial opportunity for free-market production and distribution since the second tier is left entirely up to free choice and private insurance.

A health care system that finds pockets of support from each of these accounts could also turn out to be the fairest approach to democratic reform of a system of health-care distribution.[68] We do not now have—and are not likely ever to have—a single viable theory of justice, and therefore compromises of this sort should be attractive as an approach to public policy.

The decent-minimum proposal is attractive theoretically, but it will be difficult to specify in social policy and to implement politically. The plan raises questions about whether society can fairly, consistently, and unambiguously devise a public policy that recognizes a right to health care for primary needs without creating a right to expansive and expensive forms of treatment such as liver transplantation that reduce the resources that could be put to broader use elsewhere. Nonetheless, in light of the current flux in national health systems, constructing viable systems is the major task confronting the ethics of health policy in many, and perhaps all, countries. We include in this domain the problem of *setting priorities* in the distribution and use of health resources, a problem handled in a later section of this chapter.

Fair public participation is indispensable in a process of setting the threshold of a decent minimum and in fixing the precise content of the package of goods and services to be offered (and those to be withheld). Issues of allocating, rationing, and setting priorities, as discussed later in this chapter, must be confronted in the process of public participation. When substantive standards are contested regarding a decent or sufficient level of health care, fair *procedures* for reaching agreement and implementing a social policy may be our only recourse.

Ronald Dworkin has proposed a hypothetical test of what "ideal prudent insurers" would choose.[69] He criticizes an undue use of the "rescue principle," which asserts that it is intolerable for a society to allow people to die who could have been saved by spending more money on health care. He argues that the "rescue principle" grows out of an "insulation model" that treats health care as different from and superior to all other goods. In its place,

Dworkin envisions a "prudent insurance" ideal involving "a free and unsubsidized market." This ideal market presupposes a fair distribution of wealth and income; full information about the benefits, costs, and risks of various medical procedures; and ignorance about the likelihood that any particular person will experience morbidity, either life-threatening or non-life-threatening, from diseases and accidents. Under these circumstances, whatever aggregate amount a well-informed community decides to spend on health care is just, as is the distribution pattern it chooses.

Dworkin's strategy will be difficult to implement, but it provides a good model for testing hypotheses about how much justice requires in the way of a decent minimum.

Forfeiting the Right to Health Care

If we assume that all citizens enjoy a right to a decent minimum of health care, can particular individuals forfeit the right even though they wish to retain it? The question is whether a person forfeits the right to certain forms of socially supported health care through avoidable and voluntary risky actions that result in personal ill health and that generate health care needs. Examples include patients who acquire AIDS as a result of unsafe sexual activities or intravenous drug use, smokers who develop lung cancer, workers who fail to use protective equipment in the workplace, motorcycle riders who refuse to wear helmets, and people who develop liver disease after years of excessive consumption of alcohol. Some regard it as unfair to ask all individuals who participate in insurance programs to pay higher premiums or higher taxes to support those individuals in the program who voluntarily engage in risky actions.[70] This conclusion does not conflict with the rule of fair opportunity, they argue, because risk-takers' voluntary actions reduce their opportunity.

However, can society fairly exclude risk-takers from coverage in even paradigmatic cases such as smoking? In answering this question, a society would first have to identify and differentiate the various causal factors in morbidity, such as natural causes, the social environment, and personal actions. Once these factors have been identified, solid evidence must establish that a particular disease or illness resulted from voluntary activities rather than from some other causal condition. Second, the personal actions in question must have been autonomous. If the risks are unknown at the time of action, or the action is not voluntary, individuals cannot be justly held responsible for their choices.

A fundamental problem is that it is virtually impossible to isolate causal factors in many cases of ill health because of complex causal links and limited knowledge. Medical needs often result from the conjunction of genetic predispositions, voluntary actions, effects of prior disease, and environmental and social conditions. The respective roles of these different factors are often not established, as in attempts to determine whether a particular individual's lung cancer resulted from personal cigarette smoking, passive smoking, environmental pollution, occupational conditions, or heredity (or some combination of these causal conditions).

Despite these problems, it would be fair in some circumstances to require individuals to pay higher premiums or taxes if they accept well-documented risks that may result in costly medical attention. Risk-takers could be required to contribute more to particular pools, such as insurance plans, or to pay a tax on their risky conduct, such as an increased tax on cigarettes.[71]

A more difficult question is whether it is justifiable to deny individual risk-takers equal access to scarce health care treatments that they need in part as a result of their own actions. One widely debated issue concerns patients with alcohol-related end-stage liver failure (ESLF) who need liver transplants. Donated livers are scarce, and many patients suffering from ESLF die before they can obtain transplants. In the United States in 2016, approximately 14,000 patients were awaiting a liver transplant at any one time, but only 7,841 patients received a liver transplant that year.[72] A major cause of ESLF is excessive alcohol intake that produces cirrhosis and other liver diseases. Hence, the question arises whether patients who have alcohol-related ESLF should be excluded from waiting lists for liver transplants or, if admitted to waiting lists, should be given lower priority scores. Arguments for their total exclusion or lower priority generally appeal to the risk that they will resume a pattern of alcohol abuse and again experience ESLF, thereby wasting the transplanted liver. However, studies indicate that patients with alcohol-related ESLF who receive a liver transplant and abstain from alcohol do as

well as patients whose ESLF resulted from other causes (although conditions such as a long smoking history complicate this generalization).[73] Accordingly, a good case can be made, on grounds of medical utility and fairness, for not excluding alcohol-related ESLF patients altogether and instead for requiring demonstrated and extended abstention from alcohol prior to admission to the wait list or to receiving a liver transplant.[74]

It is appropriate to exclude any patient who is consuming alcohol while seeking a liver transplant. Although it is not a national policy, most transplant centers in the United States require a six-month period of alcohol abstinence and counseling before a patient with alcohol-related ESLF can be admitted to the liver transplant waiting list. This judgment about admission to the waiting list generally requires an assessment by a multidisciplinary team that considers medical, psychological, and social factors, including those that would affect compliance, abstinence, and post-transplant success. Once on the waiting list, all patients are treated equally with other candidates for transplantation, according to scores on the Model for End-stage Liver Disease (MELD), which indicate the probability that the patient will die within three months without a transplant.

However, poignant stories depict the difficulties patients with alcohol-related liver disease face when seeking to be admitted to the wait list in the first place.[75] For patients with one specific alcohol-related cause of ESLF—acute alcohol-related hepatitis—survival for a six-month period of abstinence and counseling is often impossible, and transplant centers are now successfully transplanting them early—that is, sooner than six months.[76]

In a controversial proposal, Alvin Moss and Mark Siegler argue that patients with alcohol-related ESLF should automatically receive a lower priority ranking in the allocation of donated livers than patients who develop ESLF through no fault of their own.[77] Moss and Siegler appeal to fairness, fair opportunity, and utility in support of their proposal. Stressing that it is fair to hold people responsible for their decisions, they argue that it is "fairer to give a child dying of biliary atresia an opportunity for a *first* normal liver than it is to give a patient with [alcohol-related ESLF] who was born with a normal liver a *second* one."[78]

Even if it were established that alcoholism is a chronic disease for which individuals are not responsible, Moss and Siegler contend that individuals who have this disease have the responsibility to seek and use available and effective treatments to control their alcoholism and prevent late-stage complications, including liver failure. In this account it is fair—that is, just—to hold them responsible for their failure to do so by assigning them a lower priority for a liver transplant.

In our assessment, by contrast to the Moss-Siegler conclusion, all patients should be evaluated on a case-by-case basis, considering their medical need and probability of successful transplantation, rather than being excluded altogether or automatically assigned a lower priority. An individual can then receive a lower priority rating, as warranted.[79] There are clear examples of conditions under which personal responsibility should affect priorities. For instance, a transplant recipient who, through personal negligence, rather than inability to pay, does not take regular and sufficient immunosuppressant medication, causing the transplant to fail, should not be accepted for a second transplant or should receive lower priority for it.[80] By wasting the donated organ he or she received, the patient forfeits the opportunity for yet another chance at a normal organ.

Moss and Siegler also have a utilitarian concern about providing liver transplants to patients with alcohol-related liver diseases: They believe that the public will be less willing to donate livers if many go to patients with alcohol-related ESLF. This concern is not trivial, in light of the persistent shortage of donated livers and evidence that some people believe that these patients are less deserving of transplants because of their responsibility for their liver disease.[81] However, this concern about consequences should not replace or outweigh a fair allocation procedure. This problem underlines the need to educate the public about fair organ allocation, which includes rigorous assessment of the patient's likelihood of resuming harmful drinking and wasting a transplanted liver.

GLOBAL HEALTH POLICY AND THE RIGHT TO HEALTH

Some of the theories of justice examined early in this chapter could be presented either as *global theories* (principles of justice operate globally, not merely locally) or as *statist theories* (principles operate locally, not globally). A statist theory holds that normative requirements of justice apply only within the political state, whereas a global theory sees moral norms as applicable irrespective of political boundaries.[82] The capability theory and the well-being theory examined earlier are explicitly global theories of justice. Communitarianism and libertarianism are primarily statist theories in their typical forms. Utilitarian and many egalitarian theories could be fashioned as either global or statist.

The issues here concern (1) whether the territory in which theories, principles, and rules of justice operate should be restricted to independent political units such as nation-states or should be understood as applying globally and (2) whether traditional theories of justice have failed to provide the central concepts and principles that could be used to develop global theories of justice and global institutions. Reaction to Rawls's egalitarianism (and some say to statism) has guided much of this literature, but global theories are not primarily erected on Rawls's model.

Statist Theories and Global Theories

Rawls's theory and, until recently, many other approaches to justice in health care and health policy have been conceived in terms of the rules and policies of nation-states, where governments have historically formulated and implemented laws and policies that affect the distribution of opportunities and use of economic resources. Taxation and the use of money garnered from taxation are largely local matters of distributive justice, but some policies of states, including the expenditure of funds, are global. For example, a policy of expending state funds to help eliminate malaria from the world is a global policy.

In an eighteenth-century theory of justice that deeply affected Rawls's ideas about the *circumstances* of justice,[83] David Hume argued that rules of justice are inherently local, but that the *reasons why* rules of justice are needed (in effect required) in all nation-states apply globally. He also suggested that rules of impartiality in the formulation and application of rules of justice apply universally.[84] Rawls argued that there are *universal principles* of justice, even though many *specific rules* of justice such as those found in national health policies are not universal. The dominant conception in both Hume and Rawls is statist, yet both see the need across all nation-states for a set of impartially formulated norms of justice suitable for those states.

The idea of a right to goods and services such as a decent minimum of *health* (by contrast to a decent minimum of *health care*) through public health measures, sanitation, supply of clean drinking water, and the like can and, many argue, should be modeled on the global order that reaches beyond national health systems. A globalized world has brought a realization that projects of protecting health and maintaining healthy conditions are often international in nature and require a justice-based restructuring of the global order to be effective. One model of global justice is found in a statement of the United Nations:

> The right of everyone to the enjoyment of the highest attainable standard of physical and mental health is a human right. … [F]or millions of people throughout the world, the full realization of th[is] right … still remains a distant goal and … especially for those living in poverty, this goal is becoming increasingly remote. … [P]hysical and mental health is a most important worldwide social goal, the realization of which requires action by many other social and economic sectors in addition to the health sector.[85]

Ethical and political theories that explicitly address questions of global justice are sometimes referred to as "cosmopolitan theories," although "global theories" now seems the preferred term. In bioethics, this approach, which has deeply influenced the authors of this volume, takes as its starting point large and often catastrophic social conditions—in particular, the devastating health consequences of famine, poverty, and epidemic disease. The theory then attempts to delineate which obligations extend across national borders to address these problems. The obligations advanced are similar to those traditionally found in moral and political theory, but now globalized in scope and application.

An early influence on global theories came from Peter Singer's utilitarian theory, as discussed in Chapter 6 (pp. 220–21, and see also the discussion of utilitarian theory in Chapter 9, pp. 388–94). One reason for Singer's influence in turning philosophers' attention in a global direction was his trenchant way of pointing to the gap between the demands of fundamental *principles* of morality, such as those we treat in this book, and the *practice* of those principles at the international level. Singer succeeded in convincing many philosophers and health professionals that despite the highly demanding nature of his moral conclusions, morality requires more of us than many had thought, especially in addressing global poverty and consequent ill health.[86]

Singer's theory, which is grounded in utilitarian beneficence, is oriented toward the obligations of agents such as individuals and government officials. By contrast, the perspective of egalitarian social justice proposes that we orient theory around the moral evaluation of social institutions and their responsibilities, legitimacy, and weaknesses. The focus is not on the morality of individual choices but on the morality of the basic structure of society within which moral choices are made. The most influential global theories attempt to use a theory of justice as a model for global institutional reform—for example, reform in the structure and commitments of the World Health Organization, the World Trade Organization, and pharmaceutical pricing and marketing.

Some defenders of a global theory, including Singer, argue that Rawls unduly restricts the scope of the theory of justice. In a consistent moral theory that embraces universal principles one would expect the theory to be applied everywhere, not merely within the boundaries of particular nation-states. If the worst-off in society are a focal point of concern, as they are in Rawls's theory, the situation of the truly worst-off—the global poor—would presumably have to be addressed. The basic structure of society lies in the scattered norms and institutions of commerce, education, and public policy that affect almost everyone, and there is no clear way to, or good reason to, separate citizens from foreigners. The criterion of national citizenship, from the point of view of global theories of justice, is morally arbitrary much as race, class, and gender are. Applying rules of justice exclusively within nation-states also will have the effect of increasing disparities in wealth and well-being rather than alleviating the fundamental problems.[87]

Global theory in bioethics is motivated by problems of poor health and inequalities that are the result of many interactive effects in society. It would be bizarre for a theory of justice to look only at the distribution of health care, while ignoring the many causes of poor health and poor delivery of care, and what can and should be done about these causes. Deprivations of education cause deprivations of health, just as ill health can make it difficult to obtain a good education. A deficiency of well-being can affect the other dimensions of well-being, and all can make for poor health. In many societies a constantly compounding body of deprivations exists.

Inequalities that result from this compounding are among the most urgent for a theory of justice to address, regardless of the nation in which they occur.[88] Inequalities are not merely a matter of bad luck or personal failings. They are often a consequence of unfair social institutions that can be restructured explicitly to reduce the inequalities. If, for example, lower-quality public schools result in woeful educational inequalities, which in turn contribute to poor diet and poor health, it is within our power to alter this situation. Rawls is right to point to the pervasive effects of these institutions and their place in the theory of justice. In some theories previously discussed, especially those of Powers and Faden, and Sen and Nussbaum, the authors sensibly argue that inequalities in health and well-being brought about by severe poverty have a moral urgency at the global level.

Beyond radical inequalities in health care, somewhere around twenty million people in the developing world die each year, including approximately eight million young children, from malnutrition and diseases that can be inexpensively prevented or treated by cheap and available means. If the reach of social justice is global, such inequalities from disadvantaging conditions would be at the top of the list of conditions to be remedied.[89] The best strategies for attacking these problems remain unclear, but we can again hold out the model of a decent minimum as the goal. As noted earlier the goal in global justice is likely to be a decent minimum standard of *health*, not merely *health care*. It would be a massive gain in global justice if all persons could have a fair opportunity to achieve a reasonably good level of health and general welfare.

Allocating, Setting Priorities, and Rationing

Policies about rights to health and health care encounter countless theoretical and practical difficulties of allocating, rationing, and setting priorities. We begin to work on these problems of justice in this section by treating basic conceptual and structural matters, with primary attention to intrastate and institutional decisions, with some attention later to issues of global justice.

Allocation

Decisions about allocation of particular goods and services can have far-reaching effects on other allocations. For example, funds allocated for medical and biological research may affect the availability of training programs for physicians. Allocation decisions commonly involve selection among desirable programs. Four distinct but interrelated types of allocation can be identified. The third and fourth are of particular importance for the discussion of health care rationing later in this chapter.

1. 1. *Partitioning the comprehensive social budget.* Every large government operates with a comprehensive budget, which includes allocations for health and for other social goods, such as housing, education, culture, defense, and recreation. Health is not our only value or goal, and expenditures for other goods inevitably compete for limited resources with health-targeted expenditures. However, if a well-off society fails to allocate sufficient funds to provide adequate public health measures and access to a decent minimum of health care, its allocation system is likely to be unjust.
2. 2. *Allocating within the health budget.* Health allocation decisions must be made from within the budget portion devoted to health-related budgets. We protect and promote health in many ways besides the provision of medical care. Health policies and programs for public health, disaster relief, poverty aid, occupational safety, environmental protection, injury prevention, consumer protection, and food and drug control are all parts of society's effort to protect and promote the health of its citizens and often citizens of other nations as well.
3. 3. *Allocating within targeted budgets.* Once society has determined its budget for sectors such as public health and health care, it still must allocate its resources within each sector by selecting projects and procedures for funding. For example, determining which categories of injury, illness, or disease should receive a priority ranking is a major part of allocating health care resources. Policymakers will examine various diseases in terms of their communicability, frequency, cost, associated pain and suffering, and impact on length of life and quality of life, among other factors. Under some circumstances, for example, it might be justified, as a matter of priority ranking, to concentrate less on fatal diseases, such as some forms of cancer, and more on widespread disabling diseases, such as arthritis.
4. 4. *Allocating scarce treatments for patients.* Because health needs and desires are virtually limitless, every health care system faces some form of scarcity, and not everyone in need of a particular form of health care can gain adequate access to it. At various times and in various places, medical resources and supplies such as penicillin, insulin, kidney dialysis, cardiac transplantation, and space in intensive care units have been allocated for specific patients or classes of patients. These decisions are more difficult when an illness is life-threatening and the scarce resource potentially lifesaving. The question can become, "Who shall live when not everyone can live?"[90]

Allocation decisions of type 3 and type 4 interact. Type 3 decisions partially determine the necessity and extent of patient selection by determining the availability and supply of a particular resource. Distress in making difficult choices through explicit decisions of type 4 sometimes leads a society to modify its allocation policies at the level of type 3 to increase the supply of a particular resource. For example, in part because of controversies about criteria of access to a limited supply of dialysis machines, the US Congress passed legislation in 1972 that provides funds to ensure near-universal access to kidney dialysis and kidney transplantation for its citizens without regard to their ability to pay.[91]

Setting Priorities

Setting priorities, both in health care and in public health, is an urgent topic about just health policy.[92] Structuring clear priorities in type 3 allocation decisions has been difficult in many countries, and costs continue

to rise dramatically as a result of several factors—in particular, insurance costs, new technologies, and longer life expectancy. The difficulty in setting priorities comes in determining what ought to be done when resources are inadequate to provide all of the health benefits that it is technically possible to provide. A classic example of the problem in health policy comes from the state of Oregon.

Lessons from the Oregon plan. Legislators and citizens in Oregon engaged in a pioneering effort to set priorities in allocating health care in order to extend health insurance coverage to uninsured state residents below the poverty line. Oregon's Basic Health Services Act became a focal point for debates about justice and setting limits in health policy in the United States, including issues about access to care, cost-effectiveness, rationing, and a decent minimum. This act attempted to put into practice in Oregon what has typically been discussed only at the level of theory. Many believed that the Oregon plan would mark the beginning of a new era in coming to grips with problems of rationing in the United States, but despite this ambitious plan's important impact and influence, many moral issues regarding rationing continue to be problems for health policy.[93]

The Oregon Health Services Commission (OHSC) was charged with producing a ranked priority list of services that would define a decent minimum of coverage by Medicaid, the state and federal program that provides funds to cover medical needs for financially impoverished citizens. The goal was to expand coverage to those below the poverty level and to fund as many top priority-ranked services as possible. In 1990, the OHSC listed 1,600 ranked medical procedures ranging "from the most important to the least important" services, based in part on data about quality of well-being after treatment and cost-effectiveness analysis.

This ranking was widely criticized as unjust and arbitrary. Critics pointed to the state's ranking of tooth-capping above appendectomies as a particularly egregious example. Later Oregon reduced the list to 709 ranked services, while abandoning cost-effectiveness analysis and broadening citizen participation. The goal became to rank items on the prioritized list by clinical effectiveness and social value. These spare categories need specificity, and much ingenuity went into these efforts in Oregon.

Within the state, there was initially a strong endorsement of the list of covered services, because the plan succeeded in expanding access. However, many procedures, such as incapacitating hernias, tonsillectomy, and adenoidectomy, fell below the cutoff line of the priority list.[94] Oregon has modified the plan over the years, with the consequence of high rates of coverage loss and disenrollment from the plan, difficulty in meeting the needs of the chronically ill, increased unmet health needs, reduced access to health care services, and financial strain.[95] Oregon's priority list also had trouble managing recurring budget shortfalls.

Just strategies for setting priorities. Even before the Oregon experiment, an influential literature on setting priorities emerged from health economics, as we discussed in Chapter 6.[96] This literature urged use of cost-effectiveness analysis (CEA), the most important version being cost-utility analysis (CUA). In this strategy, health benefits are measured in terms of anticipated health gains, and costs are measured in terms of expenditures of resources. The goal is basically utilitarian: the greatest health benefits for the money expended. Health benefits are quantified, and an attempt is made to incorporate the outcome directly into public policy by measuring the impact of interventions on both the length and quality of life.

Representatives of almost all types of theory of justice other than utilitarian-oriented accounts have been invoked to raise objections to this strategy for setting limits. Charges of discrimination against infants, the elderly, and people with disabilities (especially those with permanent incapacitation and the terminally ill), as well as uncertainties about how to judge gains in quality of life, have led many to conclude that the forms of cost analysis used in Oregon allow unjust and impermissible trade-offs in setting priorities. One major problem is whether lifesaving interventions such as heart transplantation should lose out altogether in the competition for priority if other interventions such as arthritis medication provide a greater improvement in quality of life.

To address these questions fairly, numerous decisions must be made, including whether priority should go to prevention or treatment and whether lifesaving procedures take priority over other interventions. Policymakers commonly labor to set priorities in the absence of a precise or powerful decision-making instrument and in the absence of significant systems of accountability.[97] Expenditures for treatment, rather than prevention, are far

higher in the current health care systems of most industrialized nations, and government officials might justifiably choose, for example, to concentrate on preventing heart disease rather than providing individuals with heart transplants or artificial hearts. Preventive care is often more effective and more efficient in saving lives, reducing suffering, raising levels of health, and lowering costs, but preventive care typically reduces morbidity and premature mortality for unknown "statistical lives," whereas critical interventions often concentrate on known "identifiable lives." Statistical lives are more difficult for legislators and the public to understand and appreciate than are identifiable lives, which can lead to the neglect of statistical lives.

Many societies have favored identified persons and therefore have allocated resources for critical care, but good evidence shows that public health expenditures targeted at poorer communities for preventive measures, such as prenatal care, save in expenditures for future care many times the amount spent for critical care. Accordingly, moral intuitions, complex concepts, and institutional commitments may distort our thinking about the moral dilemma of whether to allocate more to rescue persons in medical need or to allocate more to prevent persons from falling into such need.

No consensus has yet been found on a way to resolve these problems in health policy and biomedical ethics, but many discussants are now open to the use of various utility-driven strategies to generate data that the public and policymakers can weigh jointly with other considerations. Public preferences, sound arguments for various policy options, and knowledge of the literature of ethics and health policy could help replace or constrain morally objectionable trade-offs indicated by economic analysis.[98]

Perhaps the major problem, as we noted in Chapter 6 (pp. 254–56), is how to establish constraints that follow from principles of justice. For example, it is unfair and unacceptable to allow forms of cost-effective rationing that adversely affect or ignore levels of health among the most disadvantaged populations, in effect worsening their condition. This generalization is obvious, but it has proved and will continue to prove extremely difficult to implement at the national level and even more so at the global level.

Rationing

We now address the types of allocation decisions categorized as types 3 and 4 in the above subsection on "Allocation" (see pp. 300–1). Both are often discussed under the topic of *rationing* and related terms, such as *triage*.[99] The choice of terms makes a difference, because each term has a different history in which changes in meaning have occurred, and each has somewhat different implications.[100] *Rationing* originally did not suggest harshness or an emergency. It meant a form of allowance, share, or portion, as when food is divided into rations in the military. Only recently has the term *rationing* been linked to limited resources and the setting of priorities in health care budgets.

Rationing has at least three relevant meanings or types. The first is related to "denial of X to individuals from lack of resources." Used with this meaning, "rationing" sometimes carries a negative connotation, especially in public debates where it is used to condemn putatively unwarranted activities of limiting health care. In a market economy, all types of goods—including health care—are to some extent rationed in this first sense by ability to pay either directly or indirectly through insurance.

A second sense of *rationing* derives not from market limits but from social policy limits: The government determines an allowance or allotment, and individuals are denied access beyond the allotted amount. Rationing gasoline and types of food during a war is a well-known example, but national health systems that do not allow purchases of goods or insurance beyond an allotted amount are equally good examples.

Finally, in a third sense of *rationing*, an allowance or allotment is distributed equitably, but those who can afford to pay for additional goods are not denied access beyond the allotted amount. In this third sense, rationing involves elements of each of the first two forms: Public policy fixes an allowance, and those who cannot afford or cannot arrange for additional units are thereby effectively denied access beyond the allowance.

We will occasionally use "rationing" in each of the three senses, while concentrating on the third; but we start with two case studies of rationing. The initial case focuses on rationing by age and the second on rationing highly expensive treatments, using heart transplantation as an example.

Rationing by age. Policies sometimes exclude or assign a lower priority to persons in a particular age group and also sometimes provide advantages to a group such as the elderly, as in Medicare entitlements in the United States. In the United Kingdom implicit rationing policies have excluded elderly, end-stage kidney patients from kidney dialysis and transplantation because of their age or expected quality of life.[101] In another example, policies for allocating transplantable kidneys in the United States give priority to young patients by assigning them additional points in the allocation formula.

Various arguments have been proposed to justify the use of age in allocation policies. Some rest on judgments about the probability of successful treatment; these arguments are usually a matter of medical utility. For instance, age may be an indicator of the probability of surviving a major operation and also a factor in the likely success of the procedure. Judgments of the probability of success may include the length of time the recipient of an organ is expected to survive, a period usually shorter for an older patient than for a younger patient. If one criterion is quality-adjusted life-years (QALYs, as discussed in Chapter 6, pp. 254–56), younger patients will typically fare better than older patients in the allocation. An example is found in a policy of the US Organ Procurement and Transplantation Network, operated by the United Network for Organ Sharing, that focuses on predicted life years from transplantation as a criterion for allocating kidneys to patients. Critics charge that use of this assessment unfairly disadvantages older patients by reducing their opportunities to receive kidney transplants.[102]

Norman Daniels has offered an influential argument for viewing age as different from properties of race and gender for purposes of fair health care allocation.[103] He appeals to prudential individual decisions about health care from the perspective of an entire lifetime. Each age group can be conceived as representing a stage in a person's life span. The goal is to allocate resources prudently throughout all stages of life within a social system that provides a fair lifetime share of health care for each citizen. As prudent deliberators, he argues, impartial persons would choose under conditions of scarcity to distribute health care over a lifetime in a way that improves their chances of attaining at least a normal life span. Daniels maintains that an impartial person would reject a pattern that reduces their chances of reaching a normal life span but that increases their chances of living beyond a normal life span if they do become elderly. Daniels maintains that an impartial person would choose to shift resources that might otherwise be consumed in prolonging the lives of the elderly to the treatment of younger persons. This policy increases each person's chance of living at least a normal life span.

Another and related theory uses a "fair-innings" argument. It considers a person's whole lifetime experience in the context of efforts to achieve equality in the distribution of health care. This argument requires that everyone have an equal chance to get fair innings (a fair period or amount of time) up to a threshold amount, but having reached the threshold (for example, seventy years of age), one is no longer entitled to receive socially supported health care. Alan Williams, a fair-innings proponent, stresses that this conception of intergenerational equity would *require*, not merely *permit*, "greater discrimination against the elderly than would be dictated simply by efficiency objectives."[104] This fair-innings argument claims to offer a justification for denying scarce health care resources to elderly patients in competition with younger patients.

All policies of age-based rationing face challenges.[105] Such rationing runs the risk of perpetuating injustice by stereotyping the elderly, treating them as scapegoats because of increased health care costs, and creating unnecessary conflicts between generations. Elderly persons in each generation will complain that, when they were younger, they did not have access to innovative technologies that were later developed, while using their taxed money for funding; and they will claim that it would be unfair to deny them beneficial technologies now. Nonetheless, to protect the health of children and many vulnerable parties, it is likely that societies will have to set a threshold age beyond which public funding for certain specified conditions is not available. This choice may be perceived as tragic, yet it may be an entirely just and justifiable policy. Indeed, it may be unjust to adopt any other policy.

Still, even if age-based allocations of health care do not necessarily violate the fair-opportunity rule, they have often been implemented in unjust ways in many countries. These allocations are a prime indicator that societies need to take systematic, publicly engaged, and closely scrutinized approaches to decisions about rationing by age and more generally about policies of equitable access.

Rationing heart transplantation. Controversies about rationing heart transplantation began shortly after it became effective in the 1980s. The number of heart transplants performed is small (3,244 in the United States in 2017), but the cost of each one is large. The current US average billed charge per heart transplant is approximately $1,382,400 and $2,564,000 for heart-lung transplants.[106] Evolving medical and political circumstances over the years have led to alterations of policy that close one gap in equity only to open other equity issues. The process that led to the Oregon Health Act in part arose from concerns about the soaring expense of organ transplants.

Despite the high cost of coverage for heart transplants, arguments have been offered for publicly funding them. The US federal Task Force on Organ Transplantation, appointed by the US Department of Health and Human Services, recommended that "a public program should be set up to cover the costs of people who are medically eligible for organ transplants but who are not covered by private insurance, Medicare, or Medicaid and who are unable to obtain an organ transplant due to the lack of funds."[107] The task force grounded its recommendation on two arguments from justice.

The first argument emphasizes the continuity between heart and liver transplants and other forms of medical care, including kidney transplants, that are already widely accepted as part of the decent minimum of health care that should be provided in a country as well off as the United States: Heart and liver transplants are comparable to other funded or fundable procedures in terms of their effectiveness in saving lives and enhancing the quality of life. In response to the claim that heart and liver transplants are too expensive, the task force argued that any burdens created by saving public funds for health should be distributed equitably rather than imposed on particular groups of patients, such as those suffering from end-stage heart or liver failure. It would be an arbitrary and incoherent policy to exclude one lifesaving procedure while funding others of comparable lifesaving potential and cost.

The second argument for equitable access focuses on practices of organ donation and procurement. Public officials often participate in efforts to increase the supply of donated organs by appealing to all citizens and residents to donate organs. It would be unfair and perhaps exploitative to solicit people, rich and poor alike, to donate organs if the organs are then distributed on the basis of ability to pay.[108] Yet, in practice in the United States, people who do not have health insurance are approximately twenty times more likely to be postmortem organ donors than they are to receive an organ transplant.[109] Arguably, it is morally inconsistent to prohibit the sale of organs while distributing donated organs according to ability to pay, and it is morally problematic to distinguish buying an organ for transplantation from buying an organ transplant procedure when a donated organ is used in the procedure; yet policies and actions that involve buying an organ for transplantation are often morally condemned without also condemning the systems that allow buying an organ transplant procedure.

These arguments are attractive appeals to coherence, but they do not establish the conclusion that justice requires expensive health care irrespective of its cost or that it is arbitrary to use a reasonably structured system of rationing that involves tough choices in setting priorities. Once a society has achieved a fair threshold determination of funding at the decent-minimum level, it legitimately may select some procedures while excluding others even when they have equal lifesaving potential and equal cost, as long as it identifies relevant differences through a fair procedure. Substantial public participation along the way will help legitimate such determinations. In the end, we should situate recommendations about funding heart transplants and all other expensive treatments in the larger context of a just social policy of allocation that will require systematic and fair setting of priorities and limits.

Rationing Scarce Treatments to Patients

Health care professionals and policymakers often must decide who will receive an available but scarce medical resource that cannot be provided to all needy people.

We concentrate here on priority methods for selecting recipients in urgent circumstances. Two broad approaches vie for primacy: (1) a utilitarian strategy that emphasizes overall maximum benefit for both patients and society, and (2) an egalitarian strategy that emphasizes the equal worth of persons and fair opportunity. We argue that these two broad approaches can justifiably and coherently be combined in many policies and practices of distribution and allocation, often through a process of specification.

We defend a system that uses two sets of substantive standards and procedural rules for rationing scarce medical resources: (1) criteria and procedures to determine an initial qualifying pool of potential recipients, such as patients eligible for heart transplantation; and (2) criteria and procedures for final selection of the recipients of particular treatments such as a heart donated for transplantation.[110] For purposes of simplicity—and because we have previously focused on the relevance and irrelevance of ability to pay for care, whether directly or indirectly through insurance—we are not here including a so-called green screen, that is, a financial screen used to determine the eligible pool of candidates.

Screening candidates for treatment. Criteria for screening potential recipients of care fall into three basic categories: constituency, progress of science, and prospect of success.[111]

Constituency factor. The first criterion uses social rather than medical factors. It is determined by clientele boundaries, such as veterans served by medical centers established for veterans; geographic or jurisdictional boundaries, such as being citizens of a legal jurisdiction served by a publicly funded hospital; and ability to pay, such as the wealthy and the highly insured. These criteria are entirely nonmedical, and they involve moral judgments that often are partial, not impartial, as in policies that exclude noncitizens or include only veterans. Each of these clientele boundaries is sometimes acceptable, but their past use has often been dubious.

For example, the Task Force on Organ Transplantation in the United States proposed that donated organs be considered national, public resources to be distributed primarily according to both the needs of patients and the probability of their successful transplantation.[112] The task force judged that foreign nationals do not have the same moral claim on organs donated in the United States as do its own citizens and residents. Its position apparently is that citizenship and residency are morally relevant properties for determining distribution. However, the task force also determined that compassion and fairness support the admission of *some* nonresident aliens. In a split vote, it recommended that nonresident aliens comprise no more than 10% of the waiting list for deceased donor kidneys and that all patients on the waiting list, including nonresident aliens, have access to organs according to the same criteria of need, probability of success, and time on the waiting list.[113]

Progress of science. A second criterion for screening potential recipients of care, that of scientific progress, is research oriented and is relevant during an experimental phase in the development of a treatment. For example, physician-investigators may justifiably exclude patients who also suffer from other diseases that might obscure the research result. The objective is to determine whether an experimental treatment is effective and how it can be improved. This criterion rests on moral and prudential judgments about the efficient and proper use of resources. The factors used to include or to exclude patients from participation in such research can be controversial, especially if persons who potentially could benefit are excluded for reasons of scientific efficiency or persons who are unlikely to benefit are continued in a clinical trial to make results of the trial acceptable to the scientific community. However, achieving scientific validity is critically important in research, and progress of science is a relevant screening criterion.

Prospect of success. Whether a treatment is experimental or routine, the likelihood of success in treating the patient is a relevant criterion because scarce medical resources should be distributed only to patients who have a reasonable chance of benefit. Ignoring this factor is unjust if it wastes resources, as in the case of organs that can be transplanted only once. Heart transplant surgeons, for example, might be tempted to list their patients as urgent priority candidates for an available heart because the patients will soon die if they do not receive a transplant, even though some of these patients may be virtually certain to die even if they do receive the heart.

Good candidates could be passed over in this process. A classification and queuing system that permits the urgency of a situation alone to determine priority is as unjust as it is inefficient. For these reasons, accountability for judgments of urgency should be sought through oversight mechanisms that either scrutinize listing criteria before they are used or review the outcomes after their use—or both.

Selecting recipients. We turn now to standards for *final selection* of recipients of particular scarce treatments. Controversy centers on the use of judgments of medical utility and social utility and impersonal mechanisms such as lotteries and queuing.

Medical utility. We here assume the generally accepted rule that judgments about medical utility should figure into decisions to ration scarce medical resources. Differences in patients' prospects for successful treatment are relevant considerations, as is maximizing the number of lives saved.

Rationing based on medical utility does not necessarily violate the requirements of egalitarian justice. For example, judgments of medical utility and of egalitarian justice, as exemplified in the fair-innings argument discussed earlier in this chapter, converge around priority for younger patients over older ones. Other things being equal, medical utility favors younger patients because they can generally be expected to have more quality-adjusted life-years ahead of them than older patients, and, other things being equal, egalitarian justice assigns them priority because they have not yet had their "fair innings" or their chances to move through stages in the life cycle.[114]

However, both medical need and probability of success, as components of medical utility, are value-laden concepts, and uncertainty often exists about likely outcomes and about the factors that contribute to success. For example, many kidney transplant surgeons dispute the importance of having a good tissue match, because minor tissue mismatches can be managed by immunosuppressant medications that reduce the body's tendency to reject transplanted organs. Insisting on the seemingly objective criterion of tissue match in distributing kidneys also can disadvantage members of racial minorities and persons with a rare tissue type, as we saw earlier in this chapter.

Criteria of medical need and probability of success sometimes come into conflict. In intensive care units, attempts to save a patient whose need is medically urgent sometimes inappropriately consume resources that could be used to save other people who will die without those resources.[115] A rule of giving priority to the sickest patients or those with the most urgent medical needs will produce unfairness when it will lead to inefficient uses of resources. Rationing schemes that either minimize or altogether exclude considerations of medical utility are indefensible, but judgments of medical utility are not always sufficient by themselves. This problem leads to questions about whether chance and queuing may legitimately be used in limited ways when selecting recipients.

Impersonal mechanisms of chance and queuing. We began this chapter by noting the oddity and unacceptability of using a lottery to distribute all social positions. However, a lottery or another system of chance is not always odd and unacceptable.[116] If medical resources are scarce and not divisible into portions, and if no major disparities exist in medical utility for patients (particularly when selection determines life or death), then considerations of fair opportunity and equal respect may justify a lottery, randomization, or queuing—depending on which procedure is the most appropriate and feasible in the circumstances.

Similar judgments have supported the use of lotteries to determine who would gain access to new drugs available only in limited supply, either because the drugs had only recently been approved or because they remained experimental. For instance, Berlex Laboratories held a lottery to distribute Betaseron, a new genetically engineered drug that appeared to slow the deterioration caused by multiple sclerosis, and several drug companies held lotteries to distribute new classes of compounds to patients with AIDS. The symbolic value of such lotteries is morally significant if they send the message that all persons deserve and will receive an equal chance at social goods.[117] These methods also determine selections with little investment of time and financial resources and can create less stress for all involved, including patients.[118] Even bypassed candidates may feel less distress at being rejected by chance than by judgments of comparative merit.

However, some impersonal selection procedures present both theoretical and practical problems. For example, the rule "first come, first served" carries the potential for injustice. Under some conditions a patient already receiving a particular treatment has a severely limited chance of survival, virtually to the point of futility, whereas other patients who need the treatment have a far better chance of survival. The question also arises whether a rule of "first come, first served" implies that those already receiving treatment have absolute priority over those who arrive later but have either more urgent needs or better prospects of success. Such a system may eventuate in injustices and should be avoided unless demonstrated to be justified in the context in which it is used.

Intensive care units (ICUs) provide a good example. Although admission to the ICU establishes a presumption in favor of continued treatment, it does not give a person an absolute claim. In decisions in neonatal intensive care about the use of extracorporeal membrane oxygenation (ECMO), a form of cardiopulmonary bypass used to support newborns with life-threatening respiratory failure, a truly scarce resource is being provided, because it is not widely available and requires the full-time presence of well-trained personnel. Robert Truog argues, rightly in our judgment, that ECMO should be withdrawn from a newborn with a poor prognosis in favor of another with a good prognosis if the latter is far more likely to survive, requires the therapy, and cannot be safely transferred to another facility.[119] Such displacement of a child from the ICU requires justification, but it need not constitute abandonment or injustice if other forms of care are provided.

Whether queuing or chance is preferable will depend largely on practical considerations, but queuing appears to be feasible and acceptable in many health care settings, including emergency medicine, ICUs, and organ transplant lists. A complicating factor is that some people do not enter the queue or a lottery in time because of factors such as slowness in seeking help, inadequate or incompetent medical attention, delay in referral, or overt discrimination. A system is probably unfair if some people can gain an advantage in access over others because they are better educated, better connected, or better situated to pay for frequent visits to physicians.

Social utility. Although criteria of social utility are suspect and controversial, the comparative social value of potential recipients is sometimes a relevant and even decisive consideration. An analogy comes from World War II, when, according to some reports, the scarce resource of penicillin was distributed to US soldiers suffering from venereal disease rather than to those suffering from battle wounds. The rationale was military need: The soldiers suffering from venereal disease could be restored to battle more quickly.[120]

An argument in favor of social-utilitarian selection is that medical institutions and personnel are trustees of society and must consider the probable future contributions of patients. We argue just below that, in rare and exceptional cases involving persons of critical social importance, criteria of social value—narrow and specific as opposed to broad and general social utility—are appropriately overriding. However, in general we need to protect the relationship of personal care and trust that would be threatened if physicians and other health care professionals regularly looked beyond their patients' needs to society's needs.

Triage: medical utility and narrow social utility. Some have invoked the model of *triage*, a French term meaning "sorting," "picking," or "choosing." It has been applied to sorting items such as wool and coffee beans according to their quality. In the delivery of health care, triage is a process of developing and using criteria for prioritization. It has been used in war, in community disasters, and in emergency rooms where injured persons have been sorted for medical treatment according to their needs and prospects. Decisions to admit and to discharge patients from ICUs sometimes involve triage. The objective is to use available medical resources as effectively and as efficiently as possible, a utilitarian rationale.[121]

Triage decisions usually appeal to *medical* utility rather than *social* utility. For example, disaster victims are generally sorted according to medical need. Those who have major injuries and will die without immediate help, but who can be salvaged, are ranked first; those whose treatment can be delayed without immediate danger are ranked second; those with minor injuries are ranked third; and those for whom no treatment will be efficacious are ranked fourth. This priority scheme is fair and does not involve judgments about individuals' comparative social worth.

However, narrow or specific judgments of comparative social worth are both inescapable and acceptable in some situations.[122] Consider, for example, an earthquake disaster in which some injured survivors are medical personnel who suffer only minor injuries. These medical personnel can justifiably receive priority for treatments if they are needed to help others. Similarly, in an outbreak of pandemic influenza, it is justifiable to inoculate physicians and nurses first to enable them to care for others. Under such conditions, a person may receive priority for treatment on grounds of narrow social utility if and only if his or her contribution is indispensable to attaining a major social goal.

It is generally legitimate to invoke medical utility followed by the use of chance or queuing for scarce resources when medical utility is roughly equal for eligible patients. It is also sometimes legitimate to invoke narrow considerations of social utility to give priority to individuals who fill social roles that are essential in achieving a better overall societal outcome. This nexus of standards should prove to be both coherent and stable, despite the mixture of appeals to egalitarian justice and utilitarian justice.

In certain contexts, such as allocation in a public health emergency resulting from pandemic influenza or a bioterrorist attack and in the allocation of organs donated by the public, it is valuable, perhaps indispensable, to engage the public when setting the allocation criteria.[123] No set of criteria is the only acceptable set, and public trust and support are essential, particularly when the public's cooperation is needed in widespread health crises and in organ donation. In most situations, the set of allocation criteria must be generally accepted as morally justified norms of justice to secure public cooperation.

CONCLUSION

We have examined an array of approaches to justice, including six different theories of justice. Although we have often most closely examined egalitarian and utilitarian approaches, we have stayed in touch with features of all six of the theories of justice analyzed early in this chapter. We have maintained that no single theory of justice or system of distributing health care is solely sufficient for constructive reflection on health policy. However, we have noted that two recent theories are especially important for their connections to the value we place on health, public health, and health care.

The richness of our moral practices and beliefs helps explain why diverse theories of justice have received skillful defenses. Absent a clear social consensus about these theories of justice, we can expect that public policies will sometimes emphasize elements of one theory and at other times elements of another theory. We have done so ourselves in this chapter. However, the existence of these several theories does not justify the tragically piecemeal approach that many countries have taken to their health care systems.

Countries lacking a comprehensive and coherent system of health care financing and delivery are destined to continue on the trail of higher costs and larger numbers of unprotected citizens. They need to improve both utility (efficiency) and justice (fairness and equality). Although justice and utility may appear to be opposing values, and have often been presented as such in moral theory, both values are indispensable in shaping a health care system. Creating a more efficient system by cutting costs and providing appropriate incentives can conflict with the goal of universal access to health care, and justice-based goals of universal coverage also may make the system less efficient. Inevitably, trade-offs between equality and efficiency will occur in almost all social systems.

Policies of just access to health care, strategies of efficiency in health care institutions, and global needs for the reduction of health-impairing conditions dwarf in social importance every other issue considered in this book. Global justice and just national health care systems are distant goals for the many millions of individuals who encounter entrenched barriers to access to health care and to better health. Every society must ration its resources, but many societies can close gaps in fair rationing more conscientiously than they have to date.

We have supported a general moral perspective from which to approach these problems. In particular, we have proposed recognition of global rights to health and enforceable rights to a decent minimum of health care in

nation-states, while recognizing that adequately securing these rights in political states and globally is an exceedingly ambitious and difficult path to pursue even when the goals of policy are strongly supported by principles and theories of justice.

Our accounts in this chapter often recognize the legitimacy of trade-offs between efficiency and justice, a position that mirrors our insistence throughout this book on the possibility of contingent conflicts between principles such as beneficence and justice and the need for trade-offs between them when conflicts occur.

NOTES

1. 1. Jorge Luis Borges, *Labyrinths* (New York: New Directions, 1962), pp. 30–35. A lottery may be a justifiable mechanism in some systems of distribution, as we note later in this chapter.
2. 2. Compare the analysis in Samuel Fleishacker, *A Short History of Distributive Justice* (Cambridge, MA: Harvard University Press, 2005).
3. 3. On the efforts of AIDS activists, see our Chapter 6, at the beginning of the subsection "Expanded access to investigational products," including references in notes (pp. 224–27). See also Carol Levine, "Changing Views of Justice after Belmont: AIDS and the Inclusion of 'Vulnerable' Subjects," in *The Ethics of Research Involving Human Subjects: Facing the 21st Century*, ed. Harold Y. Vanderpool (Frederick, MD: University Publishing Group, 1996); and Leslie Meltzer and James F. Childress, "What Is Fair Subject Selection?" in *The Oxford Textbook of Clinical Research Ethics*, ed. Ezekiel J. Emanuel et al. (New York: Oxford University Press, 2008), pp. 377–85.
4. 4. See further Guido Calabresi and Philip Bobbitt, *Tragic Choices* (New York: Norton, 1978).
5. 5. *International Union ... UAW v. Johnson Controls*, 499 U.S. 187 (1991) 111 S.Ct. 1196.
6. 6. John Stuart Mill, *Utilitarianism*, in vol. 10 of the *Collected Works of John Stuart Mill* (Toronto: University of Toronto Press, 1969), chap. 5.
7. 7. Locke, *Two Treatises of Government*, ed. Peter Laslett (Cambridge: Cambridge University Press, 1960), preface, bk. 1.6.67 and bk. 2.7.87. On classical libertarianism generally, see Eric Mack and Gerald F. Gaus, "Classical Liberalism and Libertarianism: The Liberty Tradition," in *Handbook of Political Theory*, ed. Gaus and Chandran Kukathas (London: Sage, 2004), pp. 115–30.
8. 8. Robert Nozick, *Anarchy, State, and Utopia* (New York: Basic Books, 1974), esp. pp. 149–82.
9. 9. Rawls, "Kantian Constructivism in Moral Theory" (The Dewey Lectures), *Journal of Philosophy* 77 (1980): 519.
10. 10. Rawls, *A Theory of Justice* (Cambridge, MA: Harvard University Press, 1971; rev. ed., 1999), pp. 60–67, 302–3 (1999: 52–58). Rawls later instructively restated, and partially reordered, these principles, giving reasons for their revision, in *Justice as Fairness: A Restatement*, ed. Erin Kelly (Cambridge, MA: Harvard University Press, 2001), pp. 42–43. This work has influenced the presentations made here. For an interesting handling of egalitarianism and the difference principle almost forty years after Rawls published his book, see G. A. Cohen, *Rescuing Justice and Equality* (Cambridge MA: Harvard University Press, 2008), especially pp. 68–86 and 151–65.
11. 11. Daniels, *Just Health: Meeting Health Needs Fairly* (New York: Cambridge University Press, 2008), especially pp. 29–63; and see also Daniels, *Just Health Care* (New York: Cambridge University Press, 1985), pp. 34–58. For subsequent investigations of relevant problems of egalitarianism, see Shlomi Segall, *Why Inequality Matters: Luck Egalitarianism, Its Meaning and Value* (Cambridge: Cambridge University Press, 2016).
12. 12. Leading so-called "communitarian" theorists avoid use of the label, including Alasdair MacIntyre, Charles Taylor, Michael Walzer, and Michael Sandel.
13. 13. See, for example, Michael Sandel, *Democracy's Discontent: America in Search of a Public Philosophy* (Cambridge, MA: Harvard University Press, 1996); Sandel, *Public Philosophy: Essays on Morality in Politics* (Cambridge, MA: Harvard University Press, 2005); Alasdair MacIntyre, *After Virtue*, 3rd ed. (Notre Dame, IN: University of Notre Dame Press, 2007); Michael Walzer, "The Communitarian Critique of Liberalism," *Political Theory* 18 (1990): 6–23; and Shlomo Avineri and Avner de-Shalit, eds., *Communitarianism and Individualism* (Oxford: Oxford University Press, 1992).

14. 14. Charles Taylor, "Atomism," in *Powers, Possessions, and Freedom*, ed. Alkis Kontos (Toronto: University of Toronto Press, 1979): 39–62; and, for his later views, Taylor, *A Secular Age* (Cambridge: Belknap Press of Harvard University Press, 2007, and paperback edition, 2018).
15. 15. Two instructive theories are Alasdair MacIntyre, *Whose Justice? Which Rationality?* (Notre Dame, IN: University of Notre Dame Press, 1988), esp. pp. 1, 390–403; and Michael Walzer, *Spheres of Justice: A Defense of Pluralism and Equality* (New York: Basic Books, 1983), esp. pp. 86–94. For other instructive communitarian theories in bioethics, see Daniel Callahan, "Individual Good and Common Good: A Communitarian Approach to Bioethics," *Perspectives in Biology and Medicine* 46 (2003): 496–507, esp. 500ff; Callahan, "Principlism and Communitarianism," *Journal of Medical Ethics* 29 (2003): 287–91; Mark G. Kuczewski, *Fragmentation and Consensus: Communitarian and Casuist Bioethics* (Washington: Georgetown University Press, 1997); and the wide-ranging "liberal communitarianism" in Ezekiel Emanuel, *The Ends of Human Life: Medical Ethics in a Liberal Polity* (Cambridge, MA: Harvard University Press, 1991).
16. 16. See James L. Nelson, "The Rights and Responsibilities of Potential Organ Donors: A Communitarian Approach," *Communitarian Position Paper* (Washington, DC: Communitarian Network, 1992); James Muyskens, "Procurement and Allocation Policies," *Mount Sinai Journal of Medicine* 56 (1989): 202–6; and Pradeep Kumar Prabhu, "Is Presumed Consent An Ethically Acceptable Way of Obtaining Organs for Transplant?" *Journal of the Intensive Care Society* (as published May 21, 2018), available at http://journals.sagepub.com/doi/full/10.1177/1751143718777171 (accessed August 4, 2018).
17. 17. Daniel Callahan, *What Kind of Life* (New York: Simon & Schuster, 1990), ch. 4, esp. pp. 105–13; and Callahan, *Setting Limits* (New York: Simon & Schuster, 1987), esp. pp. 106–14. Downplaying autonomy in the context of public health, Callahan argues for social pressure, including stigmatization and shaming, as legitimate efforts to change patterns of obesity, combined with government actions. See his "Obesity: Chasing an Elusive Epidemic," *Hastings Center Report* 43, no. 1 (January–February 2013): 34–40.
18. 18. See our discussion of Aristotelian theories of virtue and moral excellence in Chapter 2, pp. 49–52, and the modest extension of our account in Chapter 9, pp. 409–16.
19. 19. For an orientation to capability theories and their connection to theories of justice, see Ingrid Robeyns, "The Capability Approach," in *The Stanford Encyclopedia of Philosophy* (Winter 2016 Edition), ed. Edward N. Zalta, available at https://plato.stanford.edu/archives/win2016/entries/capability-approach/ (accessed July 22, 2018); Martha Nussbaum, *Women and Human Development: The Capabilities Approach* (Cambridge: Cambridge University Press, 2000); Harry Brighouse and Ingrid Robeyns, eds., *Measuring Justice: Primary Goods and Capabilities* (Cambridge: Cambridge University Press, 2010), especially the concluding essay by Amartya Sen, "The Place of Capability in a Theory of Justice"; Henry S. Richardson, "The Social Background of Capabilities for Freedoms," *Journal of Human Development* 8 (2007): 389–414; and Jennifer Prah Ruger, *Health and Social Justice* (New York: Oxford University Press, 2009).
20. 20. Rawls, *Justice as Fairness: A Restatement* (Cambridge, MA: Harvard University Press, 2001), pp. 58–61.
21. 21. Amartya Sen, *The Idea of Justice* (Cambridge, MA: Belknap Press of Harvard University Press, 2009), pp. 231–33. See also Sen, "Capability and Well-Being," in *The Quality of Life*, ed. Martha C. Nussbaum and Amartya K. Sen (Oxford: Clarendon Press, 1993), pp. 30–53; Sen, "Elements of a Theory of Human Rights," *Philosophy & Public Affairs* 32 (2004): 315–56; Sen, "Human Rights and Capabilities," *Journal of Human Development*, 6 (2005): 151–66; and Sen, *Commodities and Capabilities* (Oxford: Oxford University Press, 1999). See further Wiebke Kuklys, *Amartya Sen's Capability Theory: Theoretical Insights and Empirical Applications*, with a Foreword by Amartya Sen (Berlin: Springer, 2005).
22. 22. Nussbaum, *Frontiers of Justice: Disability, Nationality, Species Membership* (Cambridge, MA: Harvard University Press, 2006), esp. pp. 81–95, 155–223, 281–90, 346–52, 366–79; and, on her connection to Sen, Nussbaum, "Capabilities as Fundamental Entitlements: Sen and Social Justice," *Feminist Economics* 9 (2003): 33–59.
23. 23. Nussbaum, *Frontiers of Justice*, pp. 76–81, 392–401; and Nussbaum, "Human Dignity and Political Entitlements," in President's Council on Bioethics, *Human Dignity and Bioethics: Essays Commissioned by the President's Council on Bioethics* (Washington, DC: President's Council, March 2008), pp. 351, 377–78. See further Nussbaum, *Creating Capabilities* (Cambridge, MA: Harvard University Press, 2011).
24. 24. Nussbaum, *Frontiers of Justice*, p. 175.

25. 25. Nussbaum, "Human Dignity and Political Entitlements," pp. 357–59, 363.
26. 26. Nussbaum, "The Capabilities Approach and Animal Entitlements," in *Oxford Handbook of Animal Ethics*, ed. Tom L. Beauchamp and R. G. Frey (New York: Oxford University Press, 2011), pp. 237–38. See also her earlier statement in *Frontiers of Justice*, chap. 6.
27. 27. Nussbaum, *Frontiers of Justice*, pp. 2, 155, 166. Sridhar Venkatapuram presents another demanding capability theory dealing with important problems of bioethics. In his account, all humans have a *right* to "the capability to be healthy." *Health Justice* (Cambridge: Polity Press, 2011); for the general theory in this book, see the "Introduction" (pp. 1–38) and "Conclusion" (pp. 233–38).
28. 28. Madison Powers and Ruth R. Faden, *Social Justice: The Moral Foundations of Public Health and Health Policy* (New York: Oxford University Press, 2006), pp. 16–29. In their second book on justice, *Structural Injustice: Power, Advantage, and Human Rights* (New York: Oxford University Press, 2019), these authors deepen their theory of injustice and what morally must be done about it. This second book is heavily philosophical but still responsive to and inspired by everyday structural injustices and social movements; chapter 2 is a detailed extension of their theory of the core elements of well-being, including a revision of the names of the elements. We here followed their second book.
29. 29. Powers and Faden, *Social Justice*, pp. 37–41.
30. 30. The title of the first chapter in the first Powers-Faden book—and a major topic in their treatment of justice.
31. 31. Powers and Faden, *Social Justice*, esp. pp. 62–79, 90–95, 194–95.
32. 32. Rawls, *A Theory of Justice*, pp. 73–74 (1999: 63–65) (italics added).
33. 33. See Bernard Williams, "The Idea of Equality," in *Justice and Equality*, ed. Hugo Bedau (Englewood Cliffs, NJ: Prentice Hall, 1971), p. 135; Jeff McMahan, "Cognitive Disability, Misfortune, and Justice," *Philosophy & Public Affairs* 25 (1996): 3–35; and Janet Radcliffe Richards, "Equality of Opportunity," *Ratio* 10 (1997): 253–79.
34. 34. Pertinent to the United States, see Brian D. Smedley, Adrienne Y. Stith, and Alan R. Nelson, eds., for the Committee on Understanding and Eliminating Racial and Ethnic Disparities in Health Care, Institute of Medicine (now National Academy of Medicine), *Unequal Treatment: Confronting Racial and Ethnic Disparities in Health Care* (Washington, DC: National Academies Press, 2003); and Donald A. Barr, *Health Disparities in the United States: Social Class, Race, Ethnicity, and Health*, 2nd ed. (Baltimore, MD: Johns Hopkins University Press, 2014).
35. 35. Ivor L. Livingston, ed., *Praeger Handbook of Black American Health: Policies and Issues behind Disparities in Health,* 2nd ed. (Westport, CT: Praeger, 2004), 2 vols.; Kathryn S. Ratcliff, *Women and Health: Power, Technology, Inequality, and Conflict in a Gendered World* (Boston: Allyn & Bacon, 2002); and Nicole Lurie, "Health Disparities—Less Talk, More Action," *New England Journal of Medicine* 353 (August 18, 2005): 727–28.
36. 36. See Edward L. Hannan et al., "Access to Coronary Artery Bypass Surgery by Race/Ethnicity and Gender among Patients Who Are Appropriate for Surgery," *Medical Care* 37 (1999): 68–77; Ashish K. Jha et al., "Racial Trends in the Use of Major Procedures among the Elderly," *New England Journal of Medicine* 353 (August 18, 2005): 683–91; and T. M. Connolly, R. S. White, D. L. Sastow, et al., "The Disparities of Coronary Artery Bypass Grafting Surgery Outcomes by Insurance Status: A Retrospective Cohort Study, 2007–2014," *World Journal of Surgery* 42, no. 10 (October 1, 2018): 3240–49.
37. 37. R. H. Mehta et al., "Association of Hospital and Physician Characteristics and Care Processes with Racial Disparities in Procedural Outcomes among Contemporary Patients Undergoing Coronary Artery Bypass Grafting Surgery," *Circulation* 133, no. 2 (January 12, 2016): 124–30.
38. 38. Viola Vaccarino et al., "Sex and Racial Differences in the Management of Acute Myocardial Infarction, 1994 through 2002," *New England Journal of Medicine* 353 (August 18, 2005): 671–82; and Karen M. Freund et al., "Disparities by Race, Ethnicity, and Sex in Treating Acute Coronary Syndromes," *Journal of Women's Health* 21 (2012): 126–32. A review of Medicare data from 1992 to 2010 determined that disparities in treating acute myocardial infarction "have declined modestly but remain a problem, particularly with respect to patient sex." See Jasvinder A. Singh, Xin Lu, Said Ibrahim, and Peter Cram, "Trends in and Disparities for Acute Myocardial Infarction: An Analysis of Medicare Claims Data from 1992 to 2010," *BMC Medicine* 12 (2014): 190.
39. 39. "Cancer Screening—United States, 2010," *Morbidity and Mortality Weekly Report* 61 (January 27, 2012): 41–45; Ingrid J. Hall, Florence K. L. Tangka, Susan A. Sabatino, et al., "Patterns and Trends in

Cancer Screening in the United States," *Preventing Chronic Disease* 15 (July 26, 2018): 170465; and Tamryn F. Gray, Joycelyn Cudjoe, Jeanne Murphy, et al., "Disparities in Cancer Screening Practices among Minority and Underrepresented Populations," *Seminars in Oncology Nursing* 33, no. 2 (May 2017): 184–98. For data about cancer disparities, see National Cancer Institute, "Cancer Disparities," updated March 29, 2018, available at https://www.cancer.gov/about-cancer/understanding/disparities (accessed November 17, 2018).

40. 40. See A. O. Laiyemo et al., "Race and Colorectal Cancer Disparities," *Journal of the National Cancer Institute* 102 (April 21, 2010): 538–46; John Z. Ayanian, "Racial Disparities in Outcomes of Colorectal Cancer Screening: Biology or Barriers to Optimal Care?" *Journal of the National Cancer Institute* 102 (April 21, 2010): 511–13; and Venkata S. Tammana and Adeyinka O. Laiyemo, "Colorectal Cancer Disparities: Issues, Controversies and Solutions," *World Journal of Gastroenterology* 20, no. 4 (January 28, 2014): 869–76. For some positive signs of decreasing disparities regarding colorectal cancer, see Folasade P. May, Beth A. Glenn, Catherine M. Crespi, et al., "Decreasing Black-White Disparities in Colorectal Cancer Incidence and Stage at Presentation in the United States," *Cancer Epidemiology, Biomarkers & Prevention* 26, no. 5 (May 2017): 762–68.
41. 41. Salimah H. Meghani et al., "Advancing a National Agenda to Eliminate Disparities in Pain Care: Directions for Health Policy, Education, Practice, and Research," *Pain Medicine* 13 (January 2012): 5–28.
42. 42. Smedley, Stith, and Nelson, eds., *Unequal Treatment: Confronting Racial and Ethnic Disparities in Health Care*. For critical analyses of this report, see the special issue of *Perspectives in Biology and Medicine* 48, no. 1 suppl. (Winter 2005).
43. 43. Nakela Cook et al., "Racial and Gender Disparities in Implantable Cardioverter-Defibrillator Placement: Are They Due to Overuse or Underuse?" *Medical Care Research and Review* 68 (April 2011): 226–46. For a focus on underuse, see Paul L. Hess, Adrian F. Hernandez, Deepak L. Bhatt, et al., "Sex and Race/Ethnicity Differences in Implantable Cardioverter-Defibrillator Counseling and Use among Patients Hospitalized with Heart Failure: Findings from the Get With The Guidelines-Heart Failure Program," *Circulation* 134 (2016): 517–26; and Quinn Capers IV and Zarina Sharalaya, "Racial Disparities in Cardiovascular Care: A Review of Culprits and Potential Solutions," *Journal of Racial and Ethnic Health Disparities* 1, no. 3 (2014): 171–80.
44. 44. Keith Wailoo, *Pain: A Political History* (Baltimore, MD: Johns Hopkins University Press, 2014).
45. 45. Kelly M. Hoffman, Sophie Trawalter, Jordan R. Axt, and M. Norman Oliver, "Racial Bias in Pain Assessment and Treatment Recommendations, and False Beliefs about Biological Differences between Blacks and Whites," *PNAS* 113, no. 16 (2016): 4296–301.
46. 46. Alexander R. Green, Dana R. Carney, Daniel J. Pallin, et al., "Implicit Bias among Physicians and Its Prediction of Thrombolysis Decisions for Black and White Patients," *Journal of General Internal Medicine* 22, no. 9 (September 2007): 1231–38. On implicit bias among health care professionals and in the health care system and its possible impact on disparities, see Dayna Bowen Matthew, *Just Medicine: A Cure for Racial Inequality in American Health Care* (New York: New York University Press, 2015), esp. chaps. 2–7; William J. Hall, Mimi V. Chapman, Kent M. Lee, et al., "Implicit Racial/Ethnic Bias among Health Care Professionals and Its Influence on Health Care Outcomes: A Systematic Review," *American Journal of Public Health* 105, no. 12 (December 2015): e60–e76; Irene V. Blair, John F. Steiner, and Edward P. Havranek, "Unconscious (Implicit) Bias and Health Disparities: Where Do We Go from Here?" *Permanente Journal* 15, no. 2 (Spring 2011): 71–78; Erin Dehon, Nicole Weiss, Jonathan Jones et al., "A Systematic Review of the Impact of Implicit Racial Bias on Clinical Decision Making," *Academic Emergency Medicine* 24, no. 8 (2017): 895–904.
47. 47. Katrina Armstrong, Chanita Hughes-Halbert, and David A. Asch, "Patient Preferences Can Be Misleading as Explanations for Racial Disparities in Health Care," *Archives of Internal Medicine* 166 (May 8, 2006): 950–54; Norman G. Levinsky, "Quality and Equity in Dialysis and Renal Transplantation," *New England Journal of Medicine* 341 (November 25, 1999): 1691–93.
48. 48. John P. Roberts et al., "Effect of Changing the Priority for HLA Matching on the Rates and Outcomes of Kidney Transplantation in Minority Groups," *New England Journal of Medicine* 350 (February 5, 2004): 545–51, at 551. See also Jon J. van Rood, "Weighing Optimal Graft Survival through HLA Matching against the Equitable Distribution of Kidney Allografts," *New England Journal of Medicine* 350 (February 5, 2004): 535–36.

49. 49. Erin Hall et al., "Effect of Eliminating Priority Points for HLA-B Matching on Racial Disparities in Kidney Transplant Rates," *American Journal of Kidney Diseases* 58 (2011): 813–16. See also John S. Gill, "Achieving Fairness in Access to Kidney Transplant: A Work in Progress," *American Journal of Kidney Diseases* 58 (2011): 697–99.
50. 50. See David J. Taber, Mulugeta Gebregziabher, Kelly J. Hunt, et al., "Twenty Years of Evolving Trends in Racial Disparities for Adult Kidney Transplant Recipients," *Kidney International* 90, no. 4 (October 2016): 878–87; and Tanjala S. Purnell, Xun Luo, Lauren M. Kucirka, et al., "Reduced Racial Disparity in Kidney Transplant Outcomes in the United States from 1990 to 2012," *Journal of the American Society of Nephrology* 27 (2016): 2511–18.
51. 51. Winfred W. Williams and Francis L. Delmonico, "The End of Racial Disparities in Kidney Transplantation? Not So Fast!," *Journal of the American Society of Nephrology* 27, no. 8 (2016): 2224-26.
52. 52. Robert Bornholz and James J. Heckman, "Measuring Disparate Impacts and Extending Disparate Impact Doctrine to Organ Transplantation," *Perspectives in Biology and Medicine* 48, no. 1 suppl. (Winter 2005): S95–122.
53. 53. Carl Tishler and Suzanne Bartholomae, "The Recruitment of Normal Healthy Volunteers," *Journal of Clinical Pharmacology* 42 (2002): 365–75.
54. 54. Tom L. Beauchamp, Bruce Jennings, Eleanor Kinney, and Robert Levine, "Pharmaceutical Research Involving the Homeless," *Journal of Medicine and Philosophy* 27 (2002): 547–64; Toby L. Schonfeld, Joseph S. Brown, Meaghann Weniger, and Bruce Gordon, "Research Involving the Homeless," *IRB* 25 (September–October 2003): 17–20; and, more broadly on the use of at-risk persons as volunteers in research, Ezekiel J. Emanuel et al., "Quantifying the Risks of Non-Oncology Phase 1 Research in Healthy Volunteers: Meta-Analysis of Phase 1 Studies," *BMJ* 350 (2015), available at https://www.researchgate.net/publication/280151725_Quantifying_the_Risks_of_Non-Oncology_Phase_I_Research_in_Healthy_Volunteers_Meta-Analysis_of_Phase_I_Studies (accessed August 4, 2018).
55. 55. See the seventeen volumes published by the National Commission for the Protection of Human Subjects of Biomedical and Behavioral Research between 1975 and 1978 (published by the US Government Printing Office). The Commission studied various vulnerable populations and made recommendations for additions to, or reform of, federal policies in the United States; and many of these reports created or otherwise influenced public policy. The Commission's work was critical in turning attention to moral problems of using vulnerable populations in research. The Commission's publications and transcripts are all available at the Kennedy Institute of Ethics Library, Georgetown University, Washington, DC. See https://repository.library.georgetown.edu/bitstream/handle/10822/559326/A%20Guide%20to%20the%20National%20Commission%20of%20the%20Protection%20of%20Human%20Subjects%20of%20Biomedical%20and%20Behavioral%20Research.pdf;sequence=5.
56. 56. For analysis of the concept of exploitation, see Alan Wertheimer, *Exploitation* (Princeton, NJ: Princeton University Press, 1996); Matt Zwolinski and Alan Wertheimer, "Exploitation," *The Stanford Encyclopedia of Philosophy* (Summer 2017 Edition), ed. Edward N. Zalta, available at https://plato.stanford.edu/archives/sum2017/entries/exploitation/ (accessed July 24, 2018); and Jennifer S. Hawkins and Ezekiel J. Emanuel, eds., *Exploitation and Developing Countries: The Ethics of Clinical Research* (Princeton, NJ: Princeton University Press, 2008).
57. 57. See Carol Levine, Ruth R. Faden, Christine Grady, et al. (for the Consortium to Examine Clinical Research Ethics), "The Limitations of 'Vulnerability' as a Protection for Human Research Participants," *American Journal of Bioethics* 4 (2004): 44–49; Samia A. Hurst, "Vulnerability in Research and Health Care: Describing the Elephant in the Room?" *Bioethics* 22 (2008): 191–202; and Ruth Macklin, "Bioethics, Vulnerability, and Protection," *Bioethics* 17 (2003): 472–86.
58. 58. See Debra A. DeBruin, "Looking beyond the Limitations of 'Vulnerability': Reforming Safeguards in Research," *American Journal of Bioethics* 4 (2004): 76–78; and National Bioethics Advisory Commission, *Ethical and Policy Issues in Research Involving Human Participants*, vol. 1 (Bethesda, MD: Government Printing Office, 2001).
59. 59. See Sarah E. Hewlett, "Is Consent to Participate in Research Voluntary?" *Arthritis Care and Research* 9 (1996): 400–404; Robert M. Nelson, Tom L. Beauchamp, Victoria A. Miller, et al., "The Concept of *Voluntary* Consent," *American Journal of Bioethics* 11 (2011): 6–16; and Robert M. Nelson and Jon F.

Merz, "Voluntariness of Consent for Research: An Empirical and Conceptual Review," *Medical Care* 40 (2002), Suppl. V69–80.

60. 60. See Beauchamp, Jennings, Kinney, and Levine, "Pharmaceutical Research Involving the Homeless," 547–64; and Jennifer S. Hawkins and Ezekiel J. Emanuel, "Clarifying Confusions about Coercion," *Hastings Center Report* 35 (September–October 2005): 16–19.
61. 61. Common Rule for the Protection of Human Subjects, US Code of Federal Regulations, 45 CFR 46.116(a)(2) (implementation, January 21, 2019) https://www.ecfr.gov/cgi-bin/retrieveECFR?gp=&SID=83cd09e1c0f5c6937cd9d7513160fc3f&pitd=20180719&n=pt45.1.46&r=PART&ty=HTML (accessed May 21, 2019); and see also Ezekiel Emanuel, "Ending Concerns about Undue Inducement," *Journal of Law, Medicine, & Ethics* 32 (2004): 100–105, esp. 101.
62. 62. The justification of monetary inducement in terms of mutual benefit is defended by Martin Wilkinson and Andrew Moore, "Inducement in Research," *Bioethics* 11 (1997), 373–89; and Wilkinson and Moore, "Inducements Revisited," *Bioethics* 13 (1999): 114–30.
63. 63. See Neal Dickert and Christine Grady, "What's the Price of a Research Subject? Approaches to Payment for Research Participation," *New England Journal of Medicine* 341 (1999): 198–203; Margaret L. Russell**,** Donna G. Moralejo, and Ellen D. Burgess, "Paying Research Subjects: Participants' Perspectives," *Journal of Medical Ethics* 26 (2000): 126–30; and David Resnick, "Research Participation and Financial Inducements," *American Journal of Bioethics* 1 (2001): 54–56.
64. 64. See Abel Valenzuela, Jr., Nik Theodore, Edwin Meléndez, and Ana Luz Gonzalez, "On the Corner: Day Labor in the United States" (UCLA Center for the Study of Urban Poverty, January 2006), available at http://portlandvoz.org/wp-content/uploads/images/2009/04/national-study.pdf (accessed July 25, 2018).
65. 65. United States Census Bureau, "Health Insurance Coverage in the United States: 2016," available at https://www.census.gov/library/publications/2017/demo/p60-260.html (accessed July 25, 2018).
66. 66. See Daniels, *Just Health Care*, chaps. 3 and 4; and *Just Health: Meeting Health Needs Fairly*, pp. 46–60. See further on the right to health care Allen Buchanan and Kristen Hessler, "Specifying the Content of the Human Right to Health Care," and Buchanan and Matthew DeCamp, "Responsibility for Global Health," both published in Allen Buchanan, *Justice and Health Care: Selected Essays* (New York: Oxford University Press, 2009), as chaps. 9–10.
67. 67. Cf. Allen Buchanan, "The Right to a Decent Minimum of Health Care," *Philosophy & Public Affairs* 13 (Winter 1984): 55–78; and Buchanan, "Health-Care Delivery and Resource Allocation," in *Medical Ethics*, 2nd ed., ed. Robert M. Veatch (Boston: Jones & Bartlett, 1997), esp. pp. 337–59. Both are reprinted in Buchanan, *Justice and Health Care*, as chaps. 1–2. See also Julian Savulescu, "Justice and Healthcare: The Right to a Decent Minimum, Not Equality of Opportunity," *American Journal of Bioethics* 1 (2001), published online December 2010, available at https://www.tandfonline.com/doi/abs/10.1162/152651601300168988?journalCode=uajb20 (accessed August 5, 2018).
68. 68. Compare Peter A. Ubel et al., "Cost-Effectiveness Analysis in a Setting of Budget Constraints—Is It Equitable?" *New England Journal of Medicine* 334 (May 2, 1996): 1174–77; and Paul T. Menzel, "Justice, Liberty, and the Choice of Health System Structure," in *Medicine and Social Justice*, 2nd ed., ed. Rosamond Rhodes, Margaret P. Battin, and Anita Silvers (New York: Oxford University Press, 2012), pp. 35–46.
69. 69. This discussion of Dworkin's position is based on his *Sovereign Virtue: The Theory and Practice of Equality* (Cambridge, MA: Harvard University Press, 2000), chap. 8.
70. 70. Robert M. Veatch, "Voluntary Risks to Health: The Ethical Issues," *JAMA: Journal of the American Medical Association* 243 (January 4, 1980): 50–55; and Justin Giovannelli, Kevin Lucia, and Sabrina Corlette, "Insurance Premium Surcharges for Smokers May Jeopardize Access to Coverage," *To the Point* (A Commonwealth Fund Publication), January 2015, available at https://www.commonwealthfund.org/blog/2015/insurance-premium-surcharges-smokers-may-jeopardize-access-coverage.
71. 71. For an instructive debate about taxing unhealthy choices, see Alexander W. Cappelen and Ole Frithjof Norheim, "Responsibility, Fairness and Rationing in Health Care," *Health Policy* 76 (2006): 312–19; and, in response, Andreas Albertsen, "Taxing Unhealthy Choices: The Complex Idea of Liberal Egalitarianism in Health," *Health Policy* 120, no. 5 (2016): 561–66. For another useful debate on personal responsibility in health policy, see Phoebe Friesen, "Personal Responsibility within Health Policy: Unethical and

Ineffective," *Journal of Medical Ethics* 44, no. 1 (2018): 53–58; and, in response, Julian Savulescu, "Golden Opportunity, Reasonable Risk and Personal Responsibility for Health," *Journal of Medical Ethics* 44, no. 1 (2018): 59–61.

72. 72. For data about liver transplants, see the Organ Procurement and Transplantation Network, National Data, available at https://optn.transplant.hrsa.gov/data/view-data-reports/national-data/# (accessed November 12, 2018). Many patients die from ESLF without seeking a transplant, being referred for evaluation, or being admitted to the transplant waiting list.
73. 73. See, for example, R. Adam, V. Karam, and V. Delvart, "Evolution of Indications and Results of Liver Transplantation in Europe: A Report from the European Liver Transplant Registry (ELTR)," *Journal of Hepatology* 57 (2012): 675–88; and Claudio Augusto Marroni, Alfeu de Medeiros Fleck, Jr., Sabrina Alves Fernandes, et al., "Liver Transplantation and Alcoholic Liver Disease: History, Controversies, and Considerations," *World Journal of Gastroenterology* 24, no. 26 (2018): 2785–805.
74. 74. Arguments opposed to the total exclusion of patients with alcohol-related liver disease from liver transplantation, based on arguments from justice, appear in Carl Cohen, Martin Benjamin, and the Ethics and Social Impact Committee of the [Michigan] Transplant and Health Policy Center, "Alcoholics and Liver Transplantation," *JAMA: Journal of the American Medical Association* 265 (March 13, 1991): 1299–301; and Alexander Zambrano, "Why Alcoholics Ought to Compete Equally for Liver Transplants," *Bioethics* 30, no. 9 (2016): 689–97.
75. 75. Alexandra Rockey Fleming, "When Drinkers Suffer Liver Disease, Should Getting a Transplant Be So Hard?" *Washington Post*, January 29, 2017, available at https://www.washingtonpost.com/national/health-science/when-drinking-ruins-someones-liver-should-they-qualify-for-a-transplant/2017/01/27/7ededff0-d1c7-11e6-9cb0-54ab630851e8_story.html?utm_term=.3b5207cef8a2 (accessed November 8, 2018).
76. 76. See B. P. Lee, N. Mehta, L. Platt, et al., "Outcomes of Early Liver Transplantation for Patients with Severe Alcoholic Hepatitis," *Gastroenterology* 155, no. 2 (2018): 422–30; A. Marot, M. Dubois, E. Trépo, et al., "Liver Transplantation for Alcoholic Hepatitis: A Systematic Review with Meta-analysis," *PLOS One* 13, no. 1 (2018): e0190823; Jessica L. Mellinger and Michael L. Volk, "Transplantation for Alcohol-related Liver Disease: Is It Fair?" *Alcohol and Alcoholism* 53, no. 2 (2018): 173–77; and Eric F. Martin, "Liver Transplantation for Alcoholic Liver Disease," *Gastroenterology & Hepatology* 14, no. 9 (September 2018): 532–35.
77. 77. Alvin H. Moss and Mark Siegler, "Should Alcoholics Compete Equally for Liver Transplantation?" *JAMA: Journal of the American Medical Association* 265 (March 13, 1991): 1295–98. See also Robert M. Veatch and Lainie F. Ross, *Transplantation Ethics*, 2nd ed. (Washington, DC: Georgetown University Press, 2015), chap. 18, "Voluntary Risks and Allocation: Does the Alcoholic Deserve a New Liver?"; Walter Glannon, "Responsibility, Alcoholism, and Liver Transplantation," *Journal of Medicine and Philosophy* 23 (1998): 31–49; and James Neuberger, Karl-Heinz Schulz, Christopher Day, et al., "Transplantation for Alcoholic Liver Disease," *Journal of Hepatology* 36 (2002): 130–37. Andreas Albertsen shows how luck egalitarianism supports taking account of personal responsibility in allocating scarce liver transplants; see his "Drinking in the Last Chance Saloon: Luck Egalitarianism, Alcohol Consumption, and the Organ Transplant Waiting List," *Medicine, Health Care and Philosophy* 19 (2016): 325–38.

    In 2016, 24% of the liver transplants in the United States went to patients with alcohol liver disease, which replaced hepatitis C virus infection as the leading condition for patients receiving liver transplant. See G. Cholankeril and A. Ahmed, "Alcoholic Liver Disease Replaces Hepatitis C Virus Infection as the Leading Indication for Liver Transplantation in the United States," *Clinical Gastroenterology and Hepatology* 16, no. 8 (2018): 1356–58. On trends over time, see also Catherine E. Kling, James D. Perkins, Robert L. Carithers, et al., "Recent Trends in Liver Transplantation for Alcoholic Liver Disease in the United States," *World Journal of Hepatology* 9, no. 36 (December 28, 2017): 1315–21.

78. 78. Italics in the original. On this distinction and egalitarian justice, see John E. Roemer, *Equality of Opportunity* (Cambridge, MA: Harvard University Press, 1998).
79. 79. Some recommend using personal responsibility in modest ways. For instance, V. Thornton proposes using it as a tie-breaker in "Who Gets the Liver Transplant? The Use of Responsibility as the Tie Breaker," *Journal of Medical Ethics* 35, no. 12 (2009): 739–42. Robert Veatch concurs, proposing that we

view personal responsibility as a "small factor" that warrants deducting "a point or two" from the patient's point total for a particular transplant. See Veatch and Ross, *Transplantation Ethics*, 2nd ed., pp. 306–20. Veatch's coauthor, Lainie Ross, rejects this view, calling instead for fair access without regard to how the patients' needs came about; see pp. 306–20. From our standpoint, it is fair to treat all patients on the liver transplant waiting list equally by balancing medical needs and probability of successful outcomes, with time on the waiting list providing a tie breaker. For an overview, see James F. Childress, "Putting Patients First in Organ Allocation: An Ethical Analysis of the U.S. Debate," *Cambridge Quarterly of Healthcare Ethics* 10, no 4 (October 2001): 365–76.

80. 80. For a discussion of issues regarding fair treatment for noncompliant patients, see Nir Eyal, "Why Treat Noncompliant Patients? Beyond the Decent Minimum Account," *Journal of Medicine and Philosophy* 36 (2011): 572–88.
81. 81. See Peter A. Ubel, Jonathan Baron, and David A. Asch, "Social Acceptability, Personal Responsibility, and Prognosis in Public Judgments about Transplant Allocation," *Bioethics* 13, no. 1 (1999): 57–68, which found that people's stated unwillingness to provide scarce transplantable organs to patients who had engaged in undesirable behaviors was explained by their view that those patients were "simply less worthy of scarce transplantable organs," rather than totally by personal responsibility for organ failure or worse probable transplant outcomes. See also Ubel, "Transplantation in Alcoholics: Separating Prognosis and Responsibility from Social Biases," *Liver Transplantation and Surgery* 3, no. 3 (1997): 1–5; and Ubel, Christopher Jepson, Jonathan Baron, et al., "Allocation of Transplantable Organs: Do People Want to Punish Patients for Causing Their Illness?, *Liver Transplantation* 7, no. 7 (2001): 600–607. Another study found less negative public attitudes toward early liver transplantation for patients with acute alcoholic hepatitis and little potential impact on probable donation. G. Stroh, T. Roseli, F. Dong, and J. Forster, "Early Liver Transplantation for Patients with Acute Alcoholic Hepatitis: Public Views and the Effects on Organ Donation," *American Journal of Transplantation* 15 (2015): 1598–604.
82. 82. On this distinction and its importance, see the well-honed controversy in the following two articles: Thomas Nagel, "The Problem of Global Justice," *Philosophy & Public Affairs* 33 (2005): 113–47; and Joshua Cohen and Charles Sabel, "Extra Rempublicam Nulla Justitia?" *Philosophy & Public Affairs* 34 (2006): 147–75, esp. 148, 150ff on the nature of statist theories. See further Onora O'Neill, *Justice across Boundaries: Whose Obligations?* (Cambridge: Cambridge University Press, 2016); Nicole Hassoun, *Globalization and Global Justice: Shrinking Distance, Expanding Obligations* (Cambridge: Cambridge University Press, 2012); and Gillian Brock, *Global Justice: A Cosmopolitan Account* (Oxford: Oxford University Press, 2009).
83. 83. As Rawls notes in *A Theory of Justice*, chap. 3, § 22. See n. 3.
84. 84. Hume, *A Treatise of Human Nature*, ed. David Fate Norton and Mary Norton (Oxford: Clarendon Press, 2007), 3.2.2; and Hume, *An Enquiry concerning the Principles of Morals*, ed. Tom L. Beauchamp (Oxford: Clarendon Press, 2000), sect. 3. Public utility is a critical moral consideration in Hume's account, but overall his theory of justice is not an instance of utilitarian moral theory.
85. 85. United Nations, Office of the High Commissioner for Human Rights, Human Rights Council, "Special Rapporteur on the Right of Everyone to the Enjoyment of the Highest Attainable Standard of Physical and Mental Health," publication undated, but available at https://www.ohchr.org/en/issues/health/pages/srrighthealthindex.aspx (accessed July 25, 2018).
86. 86. See Singer's *One World: The Ethics of Globalization* (New Haven, CT: Yale University Press, 2002), especially chap. 5. This book tries to show the narrowness of strictly nationalist views of justice, policy, and world leadership. See also Peter Unger's similarly directed, *Living High and Letting Die: Our Illusion of Innocence* (New York: Oxford University Press, 1996).
87. 87. For an introduction to global justice theories and their general principles, see Gillian Brock, "Global Justice," *The Stanford Encyclopedia of Philosophy* (Spring 2017 Edition), ed. Edward N. Zalta, available at https://plato.stanford.edu/archives/spr2017/entries/justice-global/ (accessed August 2, 2108). For globalist accounts that directly pertain to bioethics, see Powers and Faden, *Social Justice*, esp. chaps. 1 and 4–7; and the contributors to *Global Health and Global Health Ethics*, ed. Solomon Benatar and Gillian Brock (Cambridge: Cambridge University Press, 2011); Joseph Millum and Ezekiel J. Emanuel, eds., *Global Justice and Bioethics* (New York: Oxford University Press, 2012); Thomas Pogge, "Human Rights and Global Health: A Research Program," *Metaphilosophy* 36 (2005): 182–209.

88. 88. See further Powers and Faden, *Social Justice*, especially pp. 69–71, and the deepening of this account in their second book, *Structural Injustice: Power, Advantage, and Human Rights*.
89. 89. UNICEF (United Nations International Children's Emergency Fund), UNICEF for Every Child, The State of the World's Children Reports, *The State of the World's Children 2*016, available at https://www.unicef.org/sowc/ (accessed August 4, 2018); Susan Lang, "Millions of Third World Children Die Needlessly Each Year Due to Mild to Moderate Malnutrition," *Cornell Chronicle* (2003), available at http://news.cornell.edu/stories/2003/06/moderate-malnutrition-kills-millions-children-needlessly; and Claire Conway, UCSF News Center, *Poor Health: When Poverty Becomes Disease*, University of California San Francisco, January 6, 2016, available at https://www.ucsf.edu/news/2016/01/401251/poor-health (accessed August 2, 2018).
90. 90. See James F. Childress, "Who Shall Live When Not All Can Live?" *Soundings: An Interdisciplinary Journal* 53, no. 4 (Winter 1970): 339–55, reprinted in *Soundings* 96, no. 3 (2013): 237–53, published with several responses.
91. 91. For the precipitating history and later developments in the United States, including ethical issues, see C. R. Blagg, "The Early History of Dialysis for Chronic Renal Failure in the United States: A View from Seattle," *American Journal of Kidney Disease* 49 (2007): 482–96; and US Institute of Medicine (now National Academy of Medicine), Committee for the Study of the Medicare End-Stage Renal Disease Program, *Kidney Failure and the Federal Government*, ed. Richard A. Rettig and Norman G. Levinsky (Washington, DC: National Academy Press, 1991).
92. 92. See Powers and Faden, *Social Justice*, chap. 6, "Setting Limits"; Norman Daniels, "Resource Allocation and Priority Setting," in *Public Health Ethics: Cases Spanning the Globe*, ed. Drue H. Barrett, Leonard W. Ortmann, Angus Dawson, et al. (New York: Singer, 2016), chap. 3; Daniels, *Just Health: Meeting Health Needs Fairly*, pp. 313–32; and Norman Daniels and James Sabin, *Setting Limits Fairly: Can We Learn to Share Medical Resources?* (New York: Oxford University Press, 2002).
93. 93. Oregon Legislature, Senate Bill 27, Or. Rev. Stat. §§ 414.025-414.750 (March 31, 1989); and Oregon Health Services Commission, *Prioritization of Health Services: A Report to the Governor and the 73rd Oregon Legislative Assembly* (March 2005), available at https://www.oregon.gov/oha/HPA/CSI-HERC/Documents/2005-Biennial-Report-to-Governor-and-Legislature.pdf (accessed July 25, 2018). See also Lawrence Jacobs, Theodore Marmor, and Jonathan Oberlander, "The Oregon Health Plan and the Political Paradox of Rationing: What Advocates and Critics Have Claimed and What Oregon Did," *Journal of Health Politics, Policy and Law* 24 (1999): 161–80; and Daniel M. Fox and Howard M. Leichter, "Rationing Care in Oregon: The New Accountability," *Health Affairs* (Summer 1991), available at https://www.healthaffairs.org/doi/pdf/10.1377/hlthaff.10.2.7 (accessed July 25, 2018).
94. 94. Health Economics Research, Inc., for the Health Care Financing Administration, *Evolution of the Oregon Plan* (Washington, DC: NTIS No. PB98–135916 INZ, December 12, 1997, as updated January 19, 1999); Oregon Department of Administrative Services, *Assessment of the Oregon Health Plan Medicaid Demonstration* (Salem: Office for Oregon Health Plan Policy and Research, 1999); and see further Oregon Health Authority, "Background on Oregon's 1115 Medicaid Demonstration Waiver," available at https://www.oregon.gov/oha/HPA/HP-Medicaid-1115-Waiver/Pages/Background.aspx (accessed August 4, 2018).
95. 95. Jonathan Oberlander, "Health Reform Interrupted: The Unraveling of the Oregon Health Plan," *Health Affairs* 26 (January–February 2007): w96–105; Rachel Solotaroff et al., "Medicaid Programme Changes and the Chronically Ill: Early Results from a Prospective Cohort Study of the Oregon Health Plan," *Chronic Illness* 1 (2005): 191–205; Matthew J. Carlson, Jennifer DeVoe, and Bill J. Wright, "Short-Term Impacts of Coverage Loss in a Medicaid Population: Early Results from a Prospective Cohort Study of the Oregon Health Plan," *Annals of Family Medicine* 4 (2006): 391–98; and Oregon Health & Science University, Center for Health Systems Effectiveness, *Evaluation of Oregon's 2012–2017 Medicaid Waiver* (December 29, 2017), available at https://www.oregon.gov/oha/HPA/ANALYTICS/Evaluation%20docs/Summative%20Medicaid%20Waiver%20Evaluation%20-%20Final%20Report.pdf (accessed July 26, 2018). See also Ronald Stock and Bruce W. Goldberg, eds., *Health Reform Policy to Practice: Oregon's Path to a Sustainable Health System: A Study in Innovation* (London: Academic Press, 2017), esp. chap. 1, Mike Bonetto, "The Oregon Narrative: History of Health Reform," and chap. 19, Thomas Daschle and Piper Nieter Su, "Leading by Example: Why Oregon Matters in the National Health Reform Discussion."

96. 96. This subsection is indebted to Powers and Faden, *Social Justice*, chap. 6.
97. 97. See Paul T. Menzel, *Medical Costs, Moral Choices* (New Haven, CT: Yale University Press, 1983), chap. 7; Peter J. Pronovost and Ruth R. Faden, "Setting Priorities for Patient Safety," *JAMA: Journal of the American Medical Association* 302 (August 26, 2009): 890–91; Akc Bergmark, Marti G. Parker, and Mats Thorslund, "Priorities in Care and Services for Elderly People: A Path without Guidelines?" *Journal of Medical Ethics* 26 (2000): 312–18; and Jennifer L. Gibson, Douglas K. Martin, and Peter A. Singer, "Setting Priorities in Health Care Organizations: Criteria, Processes, and Parameters of Success," *BMC Health Services* Research 4 (2004), available at https://www.ncbi.nlm.nih.gov/pmc/articles/PMC518972/ (accessed July 26, 2018).
98. 98. L. B. Russell et al., "The Role of Cost-Effectiveness Analysis in Health and Medicine," *JAMA: Journal of the American Medical Association* 276 (1996): 1172–77; Dan Brock, "Ethical Issues in the Use of Cost-Effectiveness Analysis," in *Public Health, Ethics, and Equity*, ed. Sudhir Anand, Fabienne Peter, and Amartya Sen (Oxford: Oxford University Press, 2004), pp. 201–23.
99. 99. For a general introduction to moral problems of health care rationing, see Greg Bognar and Iwao Hirose, *The Ethics of Health Care Rationing: An Introduction* (New York: Routledge/Taylor & Francis, 2014). For a defense of a broad use of rationing, see Peter A. Ubel, *Pricing Life: Why It's Time for Health Care Rationing* (Cambridge, MA: MIT Press, 2000).
100. 100. See Peter A. Ubel and Susan D. Goold, "'Rationing' Health Care: Not all Definitions Are Created Equal," *Archives of Internal Medicine* 158 (1998): 209–14; and Beatrix Hoffman, *Health Care for Some: Rights and Rationing in the United States since 1*930 (Chicago: University of Chicago Press, 2012).
101. 101. Rowena Mason, "Charities Call for NHS to Stop Rationing Critical Care," *Guardian*, February 18, 2017, available at https://www.theguardian.com/society/2017/feb/18/charities-stop-nhs-rationing-critical-care (accessed July 26, 2018); John McKenzie et al., "Dialysis Decision Making in Canada, the United Kingdom, and the United States," *American Journal of Kidney Diseases* 31 (1998): 12–18; Adrian Furnham and Abigail Ofstein, "Ethical Ideology and the Allocation of Scarce Medical Resources," *British Journal of Medical Psychology* 70 (1997): 51–63; and Denis Campbell and Pamela Duncan, "Patients Suffering as Direct Result of NHS Wait-Time Failures," *Guardian* (US Edition), February 8, 2018, available at https://www.theguardian.com/society/2018/feb/08/patients-suffering-direct-result-nhs-wait-time-failures (accessed July 25, 2018).
102. 102. See Peter P. Reese, Arthur L. Caplan, Roy D. Bloom, et al., "How Should We Use Age to Ration Health Care? Lessons from the Case of Kidney Transplantation," *Journal of the American Geriatric Society* 58 (2010): 1980–86; B. E. Hippen, J. R. Thistlethwaite, Jr., and L. F. Ross, "Risk, Prognosis, and Unintended Consequences in Kidney Allocation," *New England Journal of Medicine* 364 (2011): 1285–87; Leslie P. Scheunemann and Douglas B. White, "The Ethics and Reality of Rationing in Medicine," *Chest* 140 (2011): 1625–32; and M. R. Moosa and M. Kidd, "The Dangers of Rationing Dialysis Treatment: The Dilemma Facing a Developing Country," *Kidney International* 70 (September 2006): 1107–14.
103. 103. Daniels, *Just Health: Meeting Health Needs Fairly*, pp. 171–85, esp. 177–81; and Daniels, *Am I My Parents' Keeper?* (New York: Oxford University Press, 1988). For a communitarian approach to setting limits on health care for the elderly, see Daniel Callahan, *What Kind of Life*, and *Setting Limits*.
104. 104. Alan Williams, "Intergenerational Equity: An Exploration of the 'Fair Innings' Argument," *Health Economics* 6 (1997): 117–32. See also John Harris, *The Value of Life* (London: Routledge, 1985), chap. 5; and Anthony Farrant, "The Fair Innings Argument and Increasing Life Spans," *Journal of Medical Ethics* 35 (2009): 53–56.
105. 105. For criticism, see Dan W. Brock, "Justice, Health Care, and the Elderly," *Philosophy & Public Affairs* 18 (1989): 297–312; and Michael M. Rivlin, "Why the Fair Innings Argument Is Not Persuasive," *BMC Medical Ethics* 1 (2000), available at https://bmcmedethics.biomedcentral.com/articles/10.1186/1472-6939-1-1 (accessed July 24, 2018).
106. 106. National Kidney Foundation, "Financing a Transplant," *Transplant Living, http://www.transplantliving.org/beforethetransplant/finance/costs.aspx (accessed January 26, 2012); and see also Nicholas G. Smedira, "Allocating Hearts," Journal of Thoracic and Cardiovascular Surgery* 131 (2006): 775–76.
107. 107. US Department of Health and Human Services, *Report of Task Force on Organ Transplantation, Organ Transplantation: Issues and Recommendations* (Washington, DC: DHHS, 1986), pp. 105, 111.

108. 108. By contrast to these views, see Norman Daniels, "Comment: Ability to Pay and Access to Transplantation," *Transplantation Proceedings* 21 (June 1989): 3434; and also Frances M. Kamm, "The Report of the U.S. Task Force on Organ Transplantation: Criticisms and Alternatives," *Mount Sinai Journal of Medicine* 56 (May 1989): 207–20.
109. 109. Andrew A. Herring, Steffie Woolhandler, and David U. Himmelstein, "Insurance Status of U.S. Organ Donors and Transplant Recipients: The Uninsured Give, but Rarely Receive," *International Journal of Health Services* 38, no. 4 (2008): 641–52.
110. 110. Although the distinction between these two stages is analytically useful and practical in some situations, such as organ transplantation, the two stages may collapse into one in a public health crisis or in disaster medicine.
111. 111. Originally proposed by Nicholas Rescher, "The Allocation of Exotic Medical Lifesaving Therapy," *Ethics* 79 (1969): 173–86.
112. 112. United States Task Force on Organ Transplantation, *Organ Transplantation: Issues and Recommendations: Report of the Task Force on Organ Transplantation*, Public Health Service, Health Resources and Services Administration, Office of Organ Transplantation (Washington, DC: Department of Health and Human Services, 1987). On the evolution of US organ transplant policies, see Jeffrey Prottas, *The Most Useful Gift: Altruism and the Public Policy of Organ Transplants* (San Francisco: Jossey-Bass, 1994); and David L. Weimer, *Medical Governance: Values, Expertise, and Interests in Organ Transplantation* (Washington, DC: Georgetown University Press, 2010).
113. 113. United States Task Force on Organ Transplantation, *Organ Transplantation*, p. 95. A long-time policy in the US Organ Procurement and Transplantation Network (OPTN) authorized the OPTN's Ad Hoc International Relations Committee to review any transplant center in which more than 5% of its deceased donor transplants went to recipients who were nonresident aliens. This percentage was not an absolute limit but a trigger for review. In 2012 a new policy dropped this provision and called for review of all residency and citizenship data and the preparation of an annual report to ensure accountability. For a positive evaluation of the policy's transparency and accountability, see A. K. Glazier, G. M. Danovitch, and F. L. Delmonico, "Organ Transplantation for Nonresidents of the United States: A Policy for Transparency," *American Journal of Transplantation* 14 (2014): 1740–43. Many countries are concerned about "transplant tourism."
114. 114. See, for example, Govind Persad, Alan Wertheimer, and Ezekiel J. Emanuel, "Principles for Allocation of Scarce Medical Interventions," *Lancet* 373 (January 31, 2009): 423–31. These authors argue for a "complete lives" allocation system, which gives priority to "younger people who have not yet lived a complete life." See also Douglas B. White et al., "Who Should Receive Life Support during a Public Health Emergency? Using Ethical Principles to Improve Allocation Decisions," *Annals of Internal Medicine* 150 (January 20, 2009): 132–38, which proposes an allocation strategy that builds in and balances several morally relevant considerations, including "saving the most lives, maximizing the number of 'life-years' saved, and prioritizing patients who have had the least chance to live through life's stages."
115. 115. Compare and contrast Robert M. Veatch, "The Ethics of Resource Allocation in Critical Care," *Critical Care Clinics* 2 (January 1986): 73–89; Richard Wenstone, "Resource Allocation in Critical Care," in *Ethics in Anaesthesia and Intensive Care*, ed. Heather Draper and Wendy E. Scott (Oxford: Butterworth-Heinemann, 2003), pp. 145–62; Gerald R. Winslow, *Triage and Justice: The Ethics of Rationing Life-Saving Medical Resources* (Berkeley: University of California Press, 1982); and John Kilner, *Who Lives? Who Dies? Ethical Criteria in Patient Selection* (New Haven, CT: Yale University Press, 1990).
116. 116. See Duff R. Waring, *Medical Benefit and the Human Lottery* (Dordrecht, Netherlands: Springer, 2004). For broader examinations of lotteries in decision making, with particular attention to justice, see Barbara Goodwin, *Justice by Lottery* (Chicago: University of Chicago Press, 1992); and Peter Stone, *The Luck of the Draw: The Role of Lotteries in Decision Making* (New York: Oxford University Press, 2011).
117. 117. See the comment by Evan DeRenzo in Diane Naughton, "Drug Lotteries Raise Questions: Some Experts Say System of Distribution May Be Unfair," *Washington Post*, Health Section, September 26, 1995, pp. 14–15; and Childress, "Who Shall Live When Not All Can Live?"
118. 118. In Seattle, members of a closely watched committee that selected patients for dialysis when it was in limited supply felt intense pressure and stress, often accompanied by feelings of guilt. See John Broome,

"Selecting People Randomly," *Ethics* 95 (1984): 38–55, at 41; and Shana Alexander, "They Decide Who Lives, Who Dies?" *Life Magazine*, November 9, 1962, pp. 102–25.

119. 119. Robert D. Truog, "Triage in the ICU," *Hastings Center Report* 22 (May–June 1992): 13–17. See also John D. Lantos and Joel Frader, "Extracorporeal Membrane Oxygenation and the Ethics of Clinical Research in Pediatrics," *New England Journal of Medicine* 323 (August 9, 1990): 409–13; and Jonathan W. Byrnes, "A New Benchmark for Pediatric Extracorporeal Membrane Oxygenation Research," *Pediatric Critical Care Medicine* 18 (November 2017): 1072–73. For discussion of the development, early use, and eventually declining use of ECMO in neonates, see John D. Lantos, *Neonatal Bioethics: The Moral Challenges of Medical Innovation* (Baltimore, MD: Johns Hopkins University Press, 2006), pp. 52–62.
120. 120. See Ramsey, *The Patient as Person* (New Haven, CT: Yale University Press, 1970), pp. 257–58. For controversy about this example, see Robert Baker and Martin Strosberg, "Triage and Equality: An Historical Reassessment of Utilitarian Analyses of Triage," *Kennedy Institute of Ethics Journal* 2 (1992): 101–23.
121. 121. Winslow, *Triage and Justice*. But contrast Baker and Strosberg, "Triage and Equality: An Historical Reassessment." On triage, see further Robert A. Gatter and John C. Moskop, "From Futility to Triage," *Journal of Medicine and Philosophy* 20 (1995): 191–205; Michael D. Christian, Charles L. Sprung, Mary A. King, et al., on behalf of the Task Force for Mass Critical Care, "Triage: Care of the Critically Ill and Injured during Pandemics and Disasters: CHEST Consensus Statement," *Chest* 146, no. 4, Supplement (October 2014): e61S–e74S: and the report of a task force for the Society of Critical Care Medicine: Joseph L. Nates (Chair), Mark Nunnally, Ruth Kleinpell, et al., "ICU Admission, Discharge, and Triage Guidelines: A Framework to Enhance Clinical Operations, Development of Institutional Policies, and Further Research," *Critical Care Medicine* 44, no. 8 (August 2016): 1553–1602.
122. 122. See James F. Childress, "Triage in Response to a Bioterrorist Attack," in *In the Wake of Terror: Medicine and Morality in a Time of Crisis*, ed. Jonathan D. Moreno (Cambridge, MA: MIT Press, 2003), pp. 77–93.
123. 123. See T. M. Bailey, C. Haines, R. J. Rosychuk et al., "Public Engagement on Ethical Principles in Allocating Scarce Resources during an Influenza Pandemic," *Vaccine* 29 (2011): 3111–17. For a pilot effort using a deliberative democracy method to solicit community values for allocation during disasters, see Elizabeth L. Daugherty Biddison, Howard Gwon, Monica Schoch-Spana, et al., "The Community Speaks: Understanding Ethical Values in Allocation of Scarce Lifesaving Resources during Disasters," *Annals of the American Thoracic Society* 11, no. 5 (2014): 777–83; and Biddison, Gwon, Schoch-Spana, et al., "Scarce Resource Allocation during Disasters: A Mixed-Method Community Engagement Study," *Chest* 153, no. 1 (January 2018): 187–95.

8

Professional–Patient Relationships

The previous four chapters identify moral principles that underlie many judgments in biomedical ethics. This chapter uses these principles as well as the virtues discussed in Chapter 2 to interpret and specify rules and virtues of veracity, privacy, confidentiality, and fidelity, with particular attention to relationships in clinical practice, research involving human subjects, and public health.[1] The rules and virtues discussed in this chapter provide additional content to the principles and virtues elsewhere examined.

VERACITY

Codes of medical ethics have traditionally ignored obligations and virtues of veracity. The Hippocratic oath does not recommend veracity, nor does the Declaration of Geneva of the World Medical Association. The introduction to the original 1847 Code of Medical Ethics of the American Medical Association (AMA) offers flowery praise of veracity, as "a jewel of inestimable value in medical description and narrative," but the code does not mention an obligation or virtue of veracity and allows physicians virtually unlimited discretion about what to divulge to patients. The AMA's 1980 Principles of Medical Ethics recommends, without elaboration, that physicians "deal honestly with patients and colleagues," and the current version requires physicians to "be honest in all professional interactions."[2]

Despite the rather brief mention of obligations of veracity in the historical documents of medical ethics, the virtues of honesty, truthfulness, and candor are among the often deservedly praised character traits of health professionals and researchers. Nevertheless, as Annette Baier notes, honesty "is not just a hard virtue to exhibit but also a hard one to design."[3] There are many conceptual disputes, and the ground and weight of norms and virtues of veracity have also long been disputed. Henry Sidgwick's nineteenth-century observation still holds: "It does not seem clearly agreed whether Veracity is an absolute and independent obligation, or a special application of some higher principle."[4] After Sidgwick, G. J. Warnock claimed that veracity is an independent principle and virtue that ranks in importance with other general norms such as those in our framework of respect for autonomy, beneficence, nonmaleficence, and justice.[5] Our view is that rules expressing obligations of veracity are specifications of one or more of these basic general principles and that various virtues of veracity are less basic than respectfulness for the autonomy of individuals.

Obligations of Veracity

Veracity in health care refers both to timely, accurate, objective, and comprehensive transmission of information and to the way the professional fosters the patient's or subject's understanding. Hence, veracity is intimately connected to respect for autonomy. However, three primary arguments support obligations of veracity, and these arguments call on more than respect for autonomy. The first argument is based on the wide-ranging respect owed to persons in many contexts including informed-consent situations, contexts of negotiating political settlements, giving interviews to a journalist, selling a product, and the like. The second argument connects to obligations of fidelity, promise-keeping, and contract.[6] When we communicate with others, we implicitly promise that we will speak truthfully and that we will not deceive listeners. By entering into a relationship in health care or research, the patient or subject effectively enters into a contract that includes a right to receive truthful information regarding diagnosis, prognosis, procedures, and the like, just as the professional gains a right to truthful disclosures from patients and subjects. The third argument is based on the role of trust in relationships between health professionals and patients and subjects: Adherence to rules of veracity is essential to the development and maintenance of trust in these relationships.[7]

Like all other obligations discussed in this volume, veracity is prima facie binding, not an absolute obligation. Careful management of medical information—including limited disclosure, staged disclosure, nondisclosure, deception, and even lying—is occasionally justified when veracity conflicts with other obligations such as those of medical beneficence. As contexts change, the moral weights of veracity and beneficence will be heavier or lighter, and no rule is available to determine that one outweighs the other when we have to decide whether to disclose or withhold information. Accordingly, it is difficult to determine the weight of various obligations of veracity outside of specific contexts.

Nevertheless, two generalizations may be tendered: (1) Certain kinds of deception that do not involve lying are usually less difficult to justify than lying, in part because they do not threaten as deeply the relationship of trust in many contexts of health care; and (2) underdisclosure and nondisclosure of information are also less difficult to justify than lying in various circumstances now to be examined.[8]

The Disclosure of Bad News to Patients

An example of these problems is intentional nondisclosure to patients of a diagnosis of cancer or a similarly serious medical condition and a prognosis of imminent death. Different cultural traditions and philosophical accounts offer different views of the circumstances under which nondisclosure or partial disclosure is justified.[9] From our standpoint, the physician's or nurse's fundamental obligation throughout the process of disclosure is to reassure the patient while engaging sympathetically with the patient's feelings and being present as a caring, knowledgeable professional. In certain contexts, some information can be delayed and spread over a period of time, and some may justifiably never be mentioned. The physician or nurse will still have a weighty obligation to attend to proper forms of disclosure. In the case of patients who might be devastated by additional bad news beyond that already received, the primary concern of doctors and nurses should not necessarily be the disclosure of all available relevant information.

However, poor judgments about what and to whom to disclose can result in the mishandling of complex situations. In a striking case, Mr. X, a fifty-four-year-old male patient, consented to surgery for probable malignancy in his thyroid gland. After the surgery, the physician informed him that the diagnosis had been confirmed and that the tumor had been successfully removed but did not inform him of the likelihood of lung metastases and death within a few months. The physician did inform Mr. X's wife, son, and daughter-in-law about this fuller diagnosis and about the prognosis for Mr. X. All parties agreed to conceal the diagnosis and prognosis from Mr. X. The physician told Mr. X only that he needed "preventive" treatment, and Mr. X consented to irradiation and chemotherapy. The physician did not inform Mr. X of the probable causes of his subsequent shortness of breath and back pain. Unaware of his impending death, Mr. X died three months later.[10] The physician and family alike made poor judgments about withholding information, but, as we will show below, limited or staged disclosure is sometimes appropriate and virtuous.

Shifts in policies of disclosure. In recent decades, a dramatic shift has occurred in many countries in physicians' stated policies of disclosure of the diagnosis of cancer to patients. In the United States, in 1961, 88% of physicians surveyed indicated that they sought to avoid disclosing a diagnosis of cancer to patients, whereas by 1979, 98% of those surveyed reported a policy of disclosure to cancer patients.[11] A notably similar, though later (1993–98), pattern of rapid change occurred in Japan.[12] In the 1979 US survey, physicians indicated that they most frequently considered the following four factors in deciding what to disclose: age (56% of respondents), a relative's wishes regarding disclosure to the patient (51%), emotional stability (47%), and intelligence (44%).

Although veracity in the disclosure of bad news—and in disclosure throughout clinical practice—has continued to increase, some oncologists remain reluctant to disclose bad news and choose to withhold certain types of information.[13] It is unfortunate that, as in the case of Mr. X (and as reported in the 1979 survey), familial preferences sometimes unduly influence clinicians' decisions about disclosure of diagnosis and prognosis to patients. Some physicians take the view that the family can help the physician determine whether the patient is autonomous and capable of receiving information about serious risk. Although well-intended and in some cases acceptable, this approach runs the risk of begging a critical question: By what right does a physician initially

disclose information to a family without the patient's acceptance of this arrangement? Families provide important care and support for many patients, but an autonomous patient has the right to veto familial involvement. Lacking careful justification, it is unethical for a physician to first disclose information to a patient's family without the patient's authorization. The best policy is to ask the patient both at the outset and as the illness progresses about the extent to which he or she wants to involve others, if at all. This generalization holds irrespective of the patient's cultural background, which can serve as a convenient, but inappropriate, excuse for bypassing the patient by turning to another party.

Arguments for noncommunication and limited or staged communication of bad news. In some medical communities, the pendulum may have swung too far from nondisclosure to disclosure, thereby viewing physician responsibility wholly in terms of patients' rights to information and the wrongness of withholding or delaying any sort of relevant information. We argue in this subsection both that patients do not have an absolute right to be told the truth and that under some conditions, physicians should not provide, particularly all at once, full information about their patient's medical circumstances. Disclosure of pertinent information is essential as part of the process of obtaining informed consent, but the legitimate and appropriate management of information in medical care goes well beyond informed consent. We maintain that health care professionals have an obligation to manage information responsibly, sometimes limiting information and staging disclosures over time. Here virtues may guide professional conduct better than obligations or rights.

Especially dangerous is the model of a one-time delivery of all relevant information by contrast to a staged delivery. A more precautious, and justifiable, approach balances all of the patient's relevant welfare interests and the patient's informational interests and rights. This process of balancing will sometimes lead to the judgment that the physician is morally justified in withholding or postponing (or both) certain types of information. Three arguments support some degree of nondisclosure, limited disclosure, staged disclosure, and the like in health care, especially when there is "bad news," but in other cases as well.

The first argument rests on what Henry Sidgwick and others have called "benevolent deception." Such deception has long been a part of medical tradition and practice. Its defenders hold that disclosure, particularly of a prognosis of death, can violate obligations of beneficence or nonmaleficence by causing the patient anxiety, by destroying the patient's hope, by retarding or erasing a therapeutic outcome, by leading the patient to consider suicide, and the like. This line of argument—"What you don't know can't hurt and may help you"—is consequence-based. One objection to this argument rests largely on the uncertainty and unreliability of predicting consequences. A second objection rests on the moral wrongness of appealing to such consequences in these circumstances. Both objections appear in Samuel Johnson's overstatement that "I deny the lawfulness of telling a lie to a sick man for fear of alarming him. You have no business with consequences; you are to tell the truth. Besides, you are not sure what effects your telling him that he is in danger may have."[14]

Staged disclosure and cautious language about prognosis can be justified in some circumstances, despite the risk to trust between clinicians and patients. Professional norms generally support the frank and direct sharing of information about *diagnosis* and about therapeutic options, but they also may function to discourage these same qualities in sharing *prognostic* information.[15] Professional norms of disclosure should incorporate the therapeutic value of hope for patients, along with the virtues of compassion, gentleness, and sensitivity, which are often morally more important than comprehensive disclosures.

Staged disclosure and cautious language are illustrated in the following case from rehabilitation medicine.[16] For close to a month, a physician in a stroke rehabilitation unit carefully managed information in his interactions with a patient who had suffered a stroke and who asked during the first session how long it would take for his arm to improve. From the beginning the doctor knew that the patient was unlikely to recover significant use of his arm, and he offered caveats and uncertainty that did not fully match what he believed or felt. He stressed the limitations of prognostication, the unpredictability of recovery, and the need to give the brain a chance to heal. The patient received these answers well at the time, apparently preferring the physician's "ambiguous statements about the future to the alternative judgment of the permanent paralysis he fears." This indefinite, but caring and supportive, exchange continued, with the physician praising the patient's progress in walking and performing daily activities, despite residual weakness. After two weeks, the patient was enthusiastic about his progress and

asked, “How about my arm?” The physician responded, “The arm may not recover as much as the leg.” Although this statement confirmed his fears, the patient still focused on his overall progress. He had a strong hope that the physician might be mistaken, since he had repeatedly stressed his inability to prognosticate in accurate detail.

Commenting on this case later, the physician noted that having been trained in the era of “patient autonomy,” he had once felt that he “should share all available information [he] could provide about prognosis as early as possible,” trying to temper unfavorable news, for instance, about arm recovery, with positive predictions of restored walking and independent living. However, he had found both that his patients hoped for a return to their earlier lives and that bad news at any early stage tended to overwhelm good news or signs of hope. Thus, he became convinced that most of his “patients were not ready for the cold hard facts the minute they arrived at the rehabilitation hospital. They needed time to come to terms with the reality of their disabilities, while simultaneously regaining lost function.” He therefore deemed staged disclosures appropriate to sustain patients’ hopes—an understandable strategy under the circumstances. In our judgment, this physician’s decision to use staged disclosures was justified and appropriate.

The second argument is that even if professionals know all relevant information, many patients are not able to understand and appreciate the scope and implications of the information provided. Communication can be complex, especially if the patient has limited capacity to understand, and sometimes, as in the following case, intentional verbal inaccuracy can be justified: Over the years, a ninety-year-old patient, who as a young man had been decorated for courageous actions in battle, had become fearful that he would develop cancer, which he understood to be a shameful, painful, and fatal disease that would spread inexorably. He was referred for an ulcer on his lip; a biopsy established the diagnosis of squamous cell carcinoma, which would require only a short course of radiotherapy to cure, without any need for surgery or admission to a hospital. The elderly patient, tears in his eyes, asked, “It’s not cancer, is it?” The physician emphatically denied that it was cancer.[17]

The physician justified his action on several grounds. First, he pointed to the patient’s deep need for “effective reassurance.” Second, he argued that it was “more truthful” to tell this patient that he did not have cancer than to tell him that he did, because it would have been impossible to inform him that he had a curable cancer “without leaving him with a false and untrue impression” given his enduring and unchangeable beliefs. Third, addressing this patient and his concerns in his own language expressed respect rather than paternalistic arrogance. Implicit in these justifications is the conviction that, because of his apparently unalterable false beliefs, this patient lacked the capacity to adequately process the diagnosis of cancer, which, for him, entailed a prognosis of death. The physician’s decision may have been warranted in light of this patient’s condition and the depth of his false beliefs, as well as the availability of effective treatment.

A third argument is that some patients, particularly the very sick and the dying, do not want to know the truth about their condition. Despite surveys in the United States that almost universally indicate that the majority of patients want to know, some physicians maintain that patients frequently indicate by various signals, if not actual words, that they do not want the truth. This judgment may be warranted in a limited range of cases, but claims about what patients genuinely want are inherently dubious when they contradict the patients’ own reports, and this third argument sets dangerous precedents for patently paternalistic actions that masquerade as respect for autonomy.

Relying heavily on the family’s judgment that the patient would not want to receive “bad news” also sets dangerous precedents. An Italian oncologist reports that she tries to tell her patients “the complete truth,” but sometimes the patient’s family asks her not to use the word “cancer.”[18] She then relies on nonverbal communication to establish truthful therapeutic relationships with patients in line with what she views as a traditionally accepted form of Italian medical beneficence. She tries to listen carefully and assess both the verbal and nonverbal interactions, while respecting the patients’ specific needs for information.

Such practices are not without their dangers, but they need not fail to respect individual autonomy, particularly if the patient authorizes the clinician’s independent disclosure to the family. The ways in which patients exercise their autonomy will reflect their self-understandings, including sociocultural expectations and religious or other

beliefs. A choice not to know can be as autonomous and as worthy of respect as a choice to know. Accordingly, a physician needs a keen sensitivity to grasp a particular patient's preferences and to respect that patient when dispensing information according to those preferences.

Nevertheless, attending to a particular patient's expressed desire for information about prognosis is often as difficult as it is delicate, and it may be unclear in the course of decision making whether a moral mistake is being made. In one case, a twenty-six-year-old mother of two young children, had an aggressive adenocarcinoma. Following radiation therapy and two different chemotherapeutic regimens, she was fragile, but stable.[19] She was on oxygen continuously and took long-acting morphine (60 mg) three times a day. Yet, she was energetic and life-affirming. She told the new hematology/oncology fellow that she had "a feeling" based on her increased hip pain and enlarged nodules that "things aren't going as well as people tell me they are" and hoped he had some new "tricks" up his sleeve. She promptly consented to a new drug after he explained its administration, its potential adverse effects, and the ways they would try to prevent those effects, as well as his "hope that we would begin to see the long-sought-for response that might begin to heal her."

On the way to the chemotherapy unit, she said that she had heard about a woman dying of leukemia who had written several stories for her children so that they would remember her. She continued, "My girlfriend said I should do the same thing for my kids, but I don't think I'm *that* far gone, am I, Doctor Dan?" Her physician reports his "stunned silence." Unprepared for the question, he was unsure how to respond in the hall of a busy clinic, hardly an ideal setting for breaking bad news. Faced with her radiant smile, he replied: "No, Lisa, I don't think you're at that point. I'm hopeful that this new treatment will work and that you will be able to spend a lot more time with your kids." "That's what I thought, Doctor Dan," she responded. "Thanks. Now on to round three." Fourteen days later, she died, without having written her stories for her children. Years later, the physician continued to hear the echo of his last words to her, wondering whether conveying a different message, with its depressing news, would have allowed her to pen a few words or poems or to record thoughts or messages that would provide her children a living memory of their dynamic, carefree mother.

Disclosure of medical errors. "Preventable adverse events are a leading cause of death" in US hospitals, according to a report from the Institute of Medicine (now the National Academy of Medicine), which claimed in 2000 that "at least 44,000, and perhaps as many as 98,000, Americans die in hospitals each year as a result of medical errors."[20] There are disputes about the classification of preventable adverse events in health care, their numbers, their causes, and possible solutions. For instance, not all preventable adverse events—whether lethal or nonlethal, in the hospital or in other settings—are the result of medical errors or mistakes.[21] One moral responsibility is to develop systems, including training programs, to reduce medical errors and other causes of preventable adverse events. A motto in the patient safety movement holds that "errors are caused by bad systems, not by bad people."[22] Nevertheless, it is important to remove professionals deficient in personal character, knowledge, or skills who make or are likely to make medical errors.

Another moral responsibility is to disclose promptly with specificity harmful medical errors to patients and their families. Adequate disclosure often does not occur and is rarely documented when it does.[23] Evasive formulations, including the use of passive voice, ambiguous language, and euphemisms, frequently mark disclosures that are documented.[24] In one national survey of physicians, across several specialties, more than one-third of the respondents did not entirely agree with disclosing to patients all serious medical errors, and almost one-fifth admitted that because of their fear of a lawsuit they sometimes failed to fully disclose mistakes to patients.[25] The disclosure of medical error is a subset of the provision of bad news, but it is more difficult to make these disclosures because clinicians or institutions caused the harms and subsequently fear malpractice suits. These fears are understandable, but nondisclosure is almost always morally indefensible. Available evidence indicates that these fears are often overblown, and some evidence shows that disclosure may be the best policy for reducing the likelihood of malpractice suits.[26]

Other reasons for nondisclosure or limited disclosure of medical errors include concerns about harming patients and damaging patient and public trust, as well as facing staff opposition. In one case, a young boy's parents took him to a medical center for treatment of a respiratory problem. After being placed in the adult intensive care

unit, he received ten times the normal dosage of muscle relaxant, and the respirator tube slipped and pumped oxygen into his stomach for several minutes. He suffered cardiac arrest and permanent brain damage. His parents accidentally overheard a conversation that mentioned the overdose. The physician involved had decided not to inform the parents of the mistake because they "had enough on [their] minds already," but, in all likelihood, the parents justifiably felt that their tragedy was compounded by the self-protective nondisclosure and duplicity of the physician, whom they had, until this point, apparently trusted.[27]

Ethical questions arise about not only *whether* to disclose but how to disclose, how much to disclose, when to disclose, and so forth. The language of "disclosure" may inappropriately suggest a one-off provision of information, but this interpretation neglects the importance of interactive conversation between patients and clinicians. In *Talking with Patients and Families about Medical Error*, Robert Truog and colleagues appropriately focus less on specific communication skills and more on the values and attitudes that should underlie conversations between patients and clinicians regarding medical errors.[28] They emphasize five core relational values in the interactions of patients and clinicians about medical errors: transparency, respect, accountability, continuity, and kindness. These values can be understood as virtues of physicians that should be operative in conversations between clinicians and patients.

For Truog and colleagues, and for us, balancing obligations is also required. In the context of medical error, and with a firm commitment to meeting patients' and families' needs while rebuilding trust, clinicians must "recognize, weigh, and balance" competing ethical considerations such as transparency and kindness in determining what to say about medical error. A sincere apology will often be obligatory, and in some institutions an early offer of compensation is required as a matter of justice, not merely as an expression of compassion or generosity.

A key question is whether *all* medical errors should be disclosed to patients. Most of the discussion in the literature and in this section focuses on *harmful* medical errors. There is widespread agreement that all medical errors or "near misses" should be reported through appropriate institutional and other mechanisms in order to ensure accountability and improve medical care. However, less agreement exists about whether nonharmful medical errors should be disclosed to patients. The boundaries of "harm," which we analyzed in Chapter 5 (pp. 158–59) as a "setback to interests,"[29] are not always clear. Some proponents of disclosure of apparently nonharmful medical errors argue that "even errors that result in no physiological harm may still induce pain, psychological harm, and anxiety for the patient."[30] The overall effects of such errors may not be known for some time. Hence, these authors recommend "immediate disclosure of all medical errors" to the patient. They argue that this disclosure should be timely, clear, and concise, with an explanation of potential outcomes, a statement that the error is otherwise being appropriately reported to authorities, an apology, and an invitation for the patient to ask questions. They further suggest that this policy can strengthen the patient-physician relationship without a major risk from lawsuits. In our view, the ethical obligation to disclose *harmful* medical errors is more firmly established than the ethical obligation to disclose *all* medical errors. Nevertheless, clinicians should start from a presumption in favor of disclosing all medical errors and then exclude those that, in their considered judgment, after close analysis, are not material to the patient's well-being or decision making.[31]

In conclusion, the wall of almost collusive silence that has commonly surrounded medical mistakes is an unjustified and troublesome feature of medical cultures—as is its connection to what we below refer to as fidelity to one's professional colleagues.[32] It is important to build what many now call a just culture in medicine.

Deception of Third-Party Payers

Vigorous efforts to contain the costs of health care have led some physicians to use and to attempt to justify deception to secure third-party coverage. A physician in obstetrics and gynecology presented the following example: A forty-year-old woman underwent a diagnostic laparoscopy for primary infertility. Because the woman's private insurance policy did not cover this procedure for this indication, the attending surgeon instructed the resident not to write anything about infertility in the operative notes and instead to stress the two

or three fine adhesions found in the pelvic area; if these pelvic adhesions were the indication for the procedure, the patient's insurance would then cover it. When the resident refused, the attending prepared the operative note.[33]

Several studies have attempted to determine the extent to which physicians use, or would be willing to use, deception on behalf of their patients. According to one study, close to 50% of the physicians surveyed admitted that they had exaggerated the severity of their patients' medical condition so that those patients would be covered for the medical care the physicians believed they needed.[34] In another survey, 54% of the physicians surveyed acknowledged that they had deceived third-party payers in order to obtain coverage benefits for their patients' medical condition, and 39% indicated that they had exaggerated the severity of patients' conditions, altered patients' diagnoses, or reported signs or symptoms that patients did not have, with the intent to help patients obtain coverage for needed care.[35] Several of these studies indicate that a significant percentage of physicians lie or otherwise compromise truthfulness. For these physicians, fidelity to patients trumps veracity, but their actions are also sometimes motivated by their self-interest in reimbursement.[36]

Other studies have used hypothetical vignettes to determine the extent to which physicians are willing to deceive or allow deception of a third-party payer to secure approval for procedures for patients. In one study, over half of the internists surveyed sanctioned the use of deception in cases in which the patients were at immediate risk and needed coronary bypass surgery or arterial revascularization.[37] A survey of physicians and the general public found that "the public was more than twice as likely as physicians to sanction deception (26% versus 11%) and half as likely to believe that physicians have adequate time to appeal coverage decisions (22% versus 59%)."[38]

Physicians often confront a tension between their roles as patient advocates and their roles within institutional structures related to third-party payers. As before, we do not maintain that deception can never be justified in such conflicts, but physicians should place a premium on seeking alternative, nondeceptive courses of action, such as formal appeals, and should challenge unduly restrictive systems.[39] The understandable temptations of deception in these systems pose threats to physician integrity, to the moral climate in organizations, and to the fair distribution of benefits and burdens in these systems. Fidelity to patients, including strong advocacy on their behalf, is in itself virtuous, but it should not cross the boundary of the disclosure of truthful clinical information to which an impartial reviewer is entitled.

To conclude, we have argued in this section on veracity that rules of truthfulness and disclosure are profoundly important in health care. Our conclusion is not merely, as one philosopher puts it, "a strong moral presumption against lying and deception when they cause harm."[40] This approach would reduce the moral significance of lying and deception to prohibitions derived from the principle of nonmaleficence and would fail to address other reasons for disclosure in health care, including respect for autonomy. Nevertheless, many clinical contexts call for good judgment that balances all relevant ethical considerations, rather than inflexible rules about the necessity of full disclosure. No a priori decision rule is available to prioritize instant and abrupt truth-telling over staged disclosure, limited disclosure, or even nondisclosure in some contexts. This perspective follows our framework of multiple prima facie principles and our discussion of justified paternalism in Chapter 6 (pp. 234–39).

Physicians today are generally aware of how tangled and difficult lines of decision making and communication can be in medical practice. They appreciate what law and morality require in the way of honest disclosures to and understanding by patients, as well as the dangers of underdisclosure. They also know that there are many ways to respect the autonomy of patients. What we should ask of physicians is an educated and caring sensitivity to the patient's informational and therapeutic needs and preferences while carefully managing the quantity and quality of information disclosed and the pace of the disclosure. Attending to a particular patient's needs and desires for information is complicated, a fact that will endure in health care.

PRIVACY

Concerns about privacy and confidentiality pervade much of medical practice, health care administration, public health, and research. Privacy became a major concern more recently than confidentiality, which has a long history in medical ethics, but we will first discuss privacy because confidentiality is arguably a way to protect privacy in certain relationships.

Privacy in Law and Legal Theory

In the 1920s, the US Supreme Court employed an expansive "liberty" interest to protect family decision making about child-rearing, education, and the like. It later adopted the term *privacy* and expanded the individual's and the family's interests in family life, child-rearing, and other areas of personal choice. *Griswold v. Connecticut* (1965), a contraception case, set a precedent that the right to privacy not only shields information from others but also protects a sphere of individual and familial decision making and action from governmental interference. The Court's decision overturned state legislation that prohibited the use or dissemination of contraceptives. The Court determined that the right to privacy protects liberty by exempting a zone of private life from public invasion.[41]

It may seem odd and possibly inappropriate to construe a right that protects individual or familial interests as one of privacy rather than liberty or autonomy, but the right to privacy in laws often encompasses rights of limited physical and informational access as well as rights of decisional freedom. Reducing this right to a right to be free to do something or a right to act autonomously can create confusion, for reasons we will now explore.

The Concept of Privacy

Some definitions of "privacy" focus on an agent's control over access to himself or herself. These definitions mistakenly confuse privacy, which is a state or condition of limited access, with an agent's control over privacy or a right to privacy, which authorizes or empowers the agent to restrict or grant access. In short, these definitions focus on either rights or powers rather than privacy itself.[42] A person can have privacy without having a right or any other form of control over access by others. Privacy exists, for example, in long-term care facilities that render patients inaccessible or in settings where others are indifferent to or uninterested in persons. Control, whether through rights or some other mechanism, is neither necessary nor sufficient for privacy as a state or condition.[43]

Anita Allen has discerningly identified four forms of privacy that involve limited access to the person: *informational privacy*, which biomedical ethics often emphasizes; *physical privacy*, which focuses on persons and their personal spaces (the latter of which is sometimes called *locational privacy*); *decisional privacy*, which concerns personal choices; and *proprietary privacy*, which highlights property interests in the human person, for example, in a person's image or biological materials.[44] We propose, in addition, a fifth form of privacy—*relational* or *associational privacy*. It includes the family and similarly intimate relations, within which individuals make decisions in conjunction with others. As these different forms of privacy suggest, definitions of privacy are too narrow if they focus entirely on limited access to *information* about a person. Privacy, as limited access, extends to bodily products and objects or materials intimately associated with the person, as well as to the person's intimate relationships with friends, lovers, spouses, physicians, and others.

In some contexts, it is desirable to provide a tighter, more specific account of "privacy," especially when developing policies about which forms of access to which aspects of persons will constitute losses and violations of privacy. We are, however, reluctant to tinker in this way with the concept to make it more serviceable for certain types of policy. Instead, we recommend that policymakers who construct privacy policies carefully specify the conditions of access that will and will not count as a loss of privacy or as a violation of the right to privacy. The policy should demarcate the zones that are considered private and not to be invaded and should identify interests that legitimately may be balanced against privacy interests. Often the focus will be on informational privacy, but the strategy we recommend applies to a broad range of privacy interests.

The value we place on a condition of limited access or nonaccess helps explain how it comes to be categorized as private.[45] Concerns about a loss of privacy may depend not only on the kind and extent of access but also on who has access, through which means, and to which aspect of the person. As Charles Fried notes, "We may not mind that a person knows a general fact about us, and yet feel our privacy invaded if he knows the details. For instance, a casual acquaintance may comfortably know that I am sick, but it would violate my privacy if he knew the nature of the illness."[46]

Justifications of the Right to Privacy

In a highly influential 1890 article "The Right to Privacy," Samuel Warren and Louis Brandeis argued that a legal right to privacy flows from fundamental rights to life, liberty, and property.[47] They derived it largely from "the right to enjoy life—the right to be let alone." But this near-vacuous right needs more content to amount to a right to privacy. In recent discussions, several justifications of the right to privacy have been proposed, three of which deserve attention here.

One approach derives the right to privacy from a cluster of other rights. Judith Thomson argues that this cluster of personal and property rights includes rights not to be looked at; not to be caused distress (e.g., by the publication of certain information); not to be harmed, hurt, or tortured (in an effort to obtain certain information, say); and so on.[48] However, her argument relies on several allegedly foundational rights that themselves have an uncertain status, such as the right not to be looked at. We are not convinced that all of these alleged rights are rights, and some of these rights may have the right to privacy as their basis, rather than the converse.

A second justification emphasizes the instrumental value of privacy and the right to privacy: Consequentialist approaches justify rules of privacy according to their instrumental value for such ends as personal development, creating and maintaining intimate social relations, and expressing personal freedom.[49] For example, privacy may be a necessary condition for intimate relationships of love, friendship, and trust.[50]

Although we build and maintain various relationships by granting some persons and denying others certain kinds of access to ourselves, we question whether the instrumental value of privacy is the primary justification of rights to privacy. The primary justification seems to be closer to the domain of the principle of respect for autonomy, which underlies a third justification.[51] We owe respect in the sense of deference to persons' autonomous wishes not to be observed, touched, intruded on, and the like. The right to authorize or decline access is the basic right. On this basis, the justification of the right to privacy parallels the justification of the right to give an informed consent that we developed in Chapter 4 (pp. 118–23).

Joel Feinberg has observed that, historically, the language of autonomy has functioned as a political metaphor for a domain or territory in which a state is sovereign. Personal autonomy carries over the idea of a region of sovereignty for the self and a right to protect it by restricting access.[52] Other common metaphors expressing privacy in the personal domain include *zones* and *spheres* of privacy that protect autonomy.

Specifying and Balancing Rules of Privacy for Public Health Surveillance

We now consider how to specify rules and rights of privacy while allowing for justified intrusions on privacy that balance privacy interests against other interests such as the public good and the progress of medical science. We use surveillance for public health purposes as a prime example.[53]

Our goal is to find the conditions under which access to a person, and to information about a person, is warranted. Surveillance generates data that can be used for epidemiological purposes to map the incidence and prevalence of disease and for taking effective actions to protect and promote public health—for instance, to impose quarantine after exposure to communicable diseases or to notify partners that a person has a sexually transmitted disease. Epidemiological data may be anonymous, but effective actions often need personal identifiers, typically names. We will concentrate on personally identified information.

Common metaphors suggest surveillance's risk to privacy: Surveillance serves as "the eyes of public health," even the "searching or prying eyes of public health," or serves as a way to keep a "finger on the pulse of the health of a community." Each metaphor implies access to individuals and to information about them, and each indicates that surveillance entails some loss of privacy. Public health surveillance sometimes infringes on the right to privacy. Rarely do individuals consent to the collection, analysis, use, storage, and transfer of personal information for public health purposes. Hence, for the most part, identifier-based surveillance in public health without individual consent differs sharply from the collection of information in clinical care and in research.[54]

In many cases the public health rationale—based on beneficence and justice in the prevention of harm to others—provides a sufficient justification for surveillance without consent. However, public health is not a single or monolithic goal, and specificity is required to determine whether, on balance, a particular public health goal will warrant the infringement of privacy rights, as is often the case for communicable diseases such as tuberculosis and sexually transmitted diseases.[55] The justification for public health surveillance hinges on the proposed use of the data—the data by themselves will not have an impact on public health—and how effective that use is likely to be. Depending on the disease being targeted, uses of the information could include partner notification, quarantine, and isolation, or case management such as directly observed therapy for tuberculosis.

New York City's program to address uncontrolled or poorly controlled diabetes raises major questions about surveillance, in part because it targets a chronic disease rather than a communicable disease. Diabetes is a major health problem, and it is the fourth leading cause of death in New York City, where approximately 9% of the population is affected, for a total of about half a million persons with diabetes. Public health officials call diabetes an "epidemic," a technically correct term but one that also has the rhetorical advantage of evoking an image that can support the expansion of public health authority and actions. Inadequately controlled diabetes causes serious health problems, such as kidney disease, heart disease, and stroke. Beyond its tremendous health burdens for patients, it has social and economic impacts, including heavy costs to the general public.

The Department of Health and Mental Hygiene in New York City initiated a program in 2006 to require laboratories with electronic reporting capabilities to report to the department the blood sugar levels of persons with diabetes in order to determine how well their diabetes is being controlled. Interventions later added to the program include regularly notifying facilities and treating providers of their patients' blood sugar levels and sending letters to patients if they are overdue for a test or if their test results indicate that their blood sugar levels are too high.[56]

New York City public health officials hoped to develop a parallel program for HIV infection but were stymied by laws protecting privacy and confidentiality. The rationale for the program arguably became stronger as evidence accumulated that anti-retroviral therapy increases the survival and improves the quality of life of infected individuals, and also significantly reduces their risk of transmitting HIV infection to others by lowering their viral load. Here therapy serves also as prevention. Surveillance data about cell counts and viral loads could provide valuable information for both clinicians and patients and could have important public health effects.[57]

Significant ethical and policy obstacles for both the diabetes surveillance program and the proposed HIV surveillance program arise from unconsented-to infringements of the right to privacy, but even in the absence of appropriate consents, the right to privacy is not absolute and must be balanced against other ethical principles and rules. As we discussed in Chapter 1 (pp. 22–24), where we proposed a model of constrained balancing, relevant factors include the importance of the goal being sought (public health or population health, along with the avoidance or reduction of social and economic burdens); whether the surveillance program would probably realize the goal; whether the infringement of privacy rights is necessary, proportionate, and the least intrusive means consistent with realizing the goal; whether adequate security measures are in place to protect personal information (which would minimize the negative effects of overriding the right to privacy); and so forth. Cautionary notes about the diabetes program are that it does not identify persons with undiagnosed diabetes or prediabetes and that it is solely informational because it involves report and notification without additional resources for prevention and treatment services.

Proponents of privacy rights emphasize that rules protecting privacy, at least within limits, can facilitate the cooperation needed for public health programs. In this regard, there are good reasons for vigorous public engagement, with all relevant stakeholders—both professionals and members of the public—involved in the development of surveillance policies. However, apprehension is still justified that public health programs such as the diabetes program, with increased scrutiny of both patients and health care providers, may wind up alienating the communities and health professionals involved, reducing their motivation to seek or provide health care services.[58] Another concern is that the New York City diabetes program represents so-called mission creep in public health and, more dramatically, may open the way for additional and more extensive registries of sensitive data, without adequate justification, thereby compromising rights of privacy.

CONFIDENTIALITY

We surrender some privacy when we grant others access to our personal information or our bodies, but we usually retain a significant level of control over information generated about us in diagnostic and therapeutic contexts and in research. For example, physicians are obligated not to grant an insurance company or a prospective employer access to information about patients unless the patients authorize its release. When others gain access to protected information without authorization, they infringe the right to confidentiality, the right to privacy, or both.

Confidentiality could be considered a branch or subset of informational privacy. It prevents redisclosure of information originally disclosed or generated within a confidential relationship, that is, a relationship in which the confider has a reasonable and legitimate expectation that the confidant will not further disclose the information to anyone without the confider's authorization.[59] The basic difference between the right to privacy and the right to confidentiality is that an infringement of a person's right to confidentiality occurs only if the person or institution to whom the information was disclosed in confidence fails to protect the information or deliberately discloses it to someone without first-party consent. By contrast, a person who, without authorization, obtains a hospital record or gains access to a computer database violates rights of privacy but does not violate rights of confidentiality. Only the person or institution that obtains information in a confidential relationship can be charged with violating rights of confidentiality.

Traditional Rules and Evolving Practices

Rules of confidentiality appear as early as the Hippocratic Oath and continue today in national and international codes. They are arguably the most widespread rules in medical ethics across time and cultures. Nevertheless, strong rules of medical secrecy have been challenged, particularly in light of public health concerns.[60] Some commentators depict traditional confidentiality rules as little more than a convenient fiction, publicly acknowledged by health care professionals and their professional organizations but widely ignored and violated in practice. We agree that the rules are today largely ceremonial unless an underlying medical culture deeply values the protection of personal health information.

Mark Siegler argues, in an influential article, that confidentiality in medicine is a decrepit concept, because what both physicians and patients have traditionally understood as medical confidentiality no longer exists. It is "compromised systematically in the course of routine medical care." To make his point graphic, Siegler presents the case of a patient who became concerned about the number of people in the hospital with apparent access to his record and threatened to leave prematurely unless the hospital would guarantee confidentiality. Upon inquiry, Siegler discovered that many more people than he suspected had responsibilities to examine the patient's chart. When he informed the patient of the number—approximately seventy-five—he assured the patient that "these people were all involved in providing or supporting his health care services." The patient retorted, "I always believed that medical confidentiality was part of doctors' code of ethics. Perhaps you should tell me just what you people mean by 'confidentiality.'"[61]

This reaction is understandable and raises questions about the seriousness of many putative institutional and professional protections. When William Behringer tested positive for HIV at a medical center in Princeton, New Jersey, where he worked as an otolaryngologist and plastic surgeon, he received numerous phone calls of sympathy within just a few hours from members of the medical staff. Within a few days, he received similar calls from his patients, and, shortly thereafter, his surgical privileges at the medical center were suspended and his practice ruined. Despite his expectation of and request for confidentiality, the medical center took no significant precautions to protect his medical records.[62]

According to one survey of patients, medical students, and house staff about expectations and practices of confidentiality, "patients expect a more rigorous standard of confidentiality than actually exists." Virtually all patients (96%) recognized the common practice of informally discussing patients' cases for second opinions. Most (69%) expected cases to be discussed openly in professional settings to receive other opinions. A majority (51%) expected cases to be discussed in professional settings simply because they were medically interesting, and half of the patients expected cases to be discussed with office nursing staff. However, they did not expect cases to be discussed in other contexts, such as in medical journals, at parties, or with spouses or friends. By contrast, house staff and medical students reported that cases are frequently discussed with physicians' spouses (57%) and at parties (70%).[63]

Threats to confidentiality emerge in many institutions with a capacity to store and disseminate confidential medical information such as medical records, drugs prescribed, medical tests administered, and reimbursement records. In occupational medicine, computer records in corporations are growing rapidly, and data in these records can be searched. If the company routinely offers medical examinations by a corporate physician, records can be computerized and merged with claims filed by an employee's private physician for reimbursement under corporate insurance policies. Many employees are concerned that this extensive, two-track medical history will be used against them if a question of continued employment arises.

It may be possible to alter current health care practices to approximate more closely the traditional ideal of confidentiality, but a gap will remain and probably will expand in many institutions because of the need for efficient access to information in medicine. In this respect, confidentiality is indeed a decrepit practice in many settings and improving the security of information through technological measures may not adequately protect all of the interests traditionally protected by rules of confidentiality. In light of some patients' weakly informed but expansive expectations, health care professionals should communicate clearly and accurately with patients about reasonable expectations for confidentiality, as well as privacy, in different health care contexts.[64]

The Nature of Medical Confidentiality

Confidentiality is present when one person discloses information to another, whether through words or other means, and the person to whom the information is disclosed pledges, implicitly or explicitly, not to divulge that information to a third party without the confider's permission. Confidential information is private and voluntarily imparted in confidence and trust. If a patient or research subject authorizes release of the information to others, then no violation of rights of confidentiality occurs, although a loss of confidentiality and privacy occurs. In an institution in which several professionals are appropriately involved in caring for the patient, sharing the patient's information could be warranted not only by medical beneficence but also by the patient's implied consent in accepting care at that institution.[65]

There exist acknowledged and justifiable exceptions or limits to the kind of information that can be considered confidential in policy and practice. For example, legal rules may set limits to confidentiality, as when they require practitioners to report gunshot wounds and venereal diseases to public health and other officials. Some unwanted disclosures of apparently confidential information to third parties may not breach confidentiality because of the context in which the information was originally gathered. For example, IBM physician Martha Nugent informed her employer of her belief that an employee, Robert Bratt, had a problem of paranoia that affected his behavior on the job.[66] Bratt knew that Nugent had been retained by IBM to examine him but expected conventional medical confidentiality. The company held that the facts disclosed by Nugent were

necessary for evaluating Bratt's request for a job transfer and, under law, were a legitimate business communication. In our view, it is a reasonable conclusion that such information is not confidential by the standards of medical confidentiality relevant to this kind of case and that Nugent was not bound by obligations of confidentiality in the same way a private physician would have been. However, as a moral—by contrast to a legal—matter, Nugent arguably should have disclosed to Bratt that conventional standards of medical confidentiality did not apply in their relationship because of her contractual obligations to IBM.

Contracts calling for limited disclosures are legitimate as long as employees are aware of, or should be aware of, provisions in the contract. A similar point applies to military physicians who have a dual responsibility to the soldier as patient and to the military. Nevertheless, the company and the military, along with the physicians in each context, have a moral responsibility to ensure that employees and soldiers understand, at the outset, the conditions under which rules of confidentiality and protections of privacy do and do not apply. This moral responsibility also extends to incarcerated persons who face complicated institutional rules and forms of infringement of confidentiality in prisons.[67] More broadly, patients and research subjects in any setting should be informed about the operative rules of confidentiality, including exceptions to those rules. This standard is particularly important in view of the public's widespread, but often incorrect, assumptions about confidential relations with health care professionals.

The Justification of Obligations of Confidentiality

Many of the goods of medicine and research could be realized without rules of confidentiality. On what basis, then, can we justify a system of extensive, often expensive and inefficient, protections of confidentiality? Two types of argument justify rules to protect confidentiality: (1) consequence-based arguments and (2) arguments from autonomy and privacy rights. These arguments also help us determine legitimate exceptions to rules of confidentiality.

Consequence-based arguments. Patients would be reluctant to disclose full and forthright information or authorize a complete examination and a full battery of tests if they could not trust physicians to conceal some information from third parties. Without such information, physicians would not be able to make accurate diagnoses and prognoses or recommend the best course of treatment.

In the precedential *Tarasoff* case, a patient told his therapist that he wanted to kill a young woman who had spurned his attention. The therapist alerted the university police but did not warn the intended victim. After the patient killed the young woman, the family brought a suit alleging that the therapist should have warned the intended victim. In this case, the California Supreme Court examined the basis and limits of confidentiality.[68] Both the majority opinion, which affirmed that therapists have an obligation to warn third parties of their patients' threatened violence, and the dissenting opinion, which denied such an obligation, used consequentialist arguments. Their debate hinged on different predictions and assessments of the consequences of a rule that *requires* therapists to breach confidentiality by warning intended victims of a client's threatened violence by contrast to a rule that *allows* therapists to breach confidentiality when a member of the public is endangered.

The majority opinion pointed to the victims who would be saved, such as the young woman who had been killed in this case, and contended that a professional's *obligation* to disclose information to third parties could be justified by the need to protect such potential victims. By contrast, the minority opinion contended that if it were common practice in such cases to override obligations of confidentiality, the fiduciary relation between the patient and the doctor would soon erode and collapse. Patients would lose confidence in psychotherapists and would refrain from disclosing information crucial to effective therapy. As a result, violent assaults would increase because dangerous persons would refuse to seek psychotherapeutic or psychiatric aid or to disclose relevant information, such as their violent fantasies. These claims about rules of confidentiality hinge significantly on empirical claims about which rule more effectively protects the interests of other persons along with the interests of patients.

In cases of other legally accepted and mandated exceptions to confidentiality—such as requirements to report contagious diseases, child abuse, and gunshot wounds—no substantial evidence exists that these requirements

have either reduced prospective patients' willingness to seek treatment and to cooperate with physicians or significantly impaired the physician-patient relationship.[69] However, one study did find that state laws that impose an enforceable legal duty to warn intended victims of a patient's threatened violence increase the rate of homicides by 5%.[70]

In a consequence-oriented framework, *nonabsolute* rules of confidentiality are attractive and acceptable as long as it is understood that when physicians or other health professionals breach confidentiality, they infringe their patients' rights. Such an infringement will almost always have negative effects for confiders. A physician who breaks confidence cannot ignore the potential for eroding the system of medical confidentiality, trust, and fidelity. In short, an acceptable consequentialist justification for breaching confidentiality must take into account all probable consequences, and policymakers must balance the probable benefits and risks of different rules of confidentiality in light of the best available evidence.

Arguments from autonomy and privacy rights. A second set of arguments to justify rules and rights of confidentiality derives from both the principle of respect for autonomy and rules of privacy. The claim is that rights of both autonomous choice and privacy jointly justify rights of confidentiality. Like the first argument, this second argument does not support absolute rules or absolute rights of confidentiality. When rules and rights of confidentiality are used as absolute shields, they can eventuate in outrageous and preventable injuries and harms.[71] (We have treated the claim that there are no absolute principles, rules, or rights in several previous chapters.)

Justified Infringements of Rules of Confidentiality

Infringements of prima facie rules and rights of confidentiality can be justified in some circumstances in which third parties face serious harms. We here concentrate on these situations, while noting that paternalistic breaches of confidentiality to remove or prevent harms to the patient are also sometimes ethically justifiable. (See further our discussion of paternalism in Chapter 6, pp. 237–39.)

Debates about when, if ever, breaches of doctor-patient confidentiality can be justified to protect the public interest intensified in bioethics and beyond after March 25, 2015, when a Germanwings co-pilot, Andreas Lubitz, appeared to intentionally crash an Airbus 320 in the French Alps, killing all 150 passengers and crew. An investigation determined that the crash was deliberately caused—indeed premeditated—by the co-pilot, who suffered from serious, long-term mental health problems, including severe depression and suicidal ideation. In addition to out-patient treatment, he underwent a short-term hospitalization for these problems, and his doctors deemed him mentally unfit to fly.[72] However, his doctors were also subject to article 9 governing doctor-patient confidentiality in the German Medical Association's professional code: "Physicians are obliged to maintain confidentiality regarding everything confided in them, or becoming known to them, in their capacity as a physician, including after the death of the patient."[73] This code recognizes valid exceptions to this rule only when patients release physicians from the obligation to maintain confidentiality or "insofar as disclosure is necessary in order to safeguard a higher-ranking legally protected interest." In the Airbus 320 case the doctors did not disclose the information to third parties who were positioned to stop their patient from flying a commercial aircraft.

Assessing and reducing risks to others. In assessing which risks to third parties outweigh rules or rights of confidentiality, the probability that a harm will materialize and the magnitude of that harm must be balanced against the moral weight of norms of confidentiality and the possible harms that might occur by breaching those norms. The chart of risk assessment introduced on p. 245 supplies the basic categories:

		Magnitude of Harm	
		Major	*Minor*
Probability of	*High*	1	2

Harm

Low 3 4

As a health professional's reasoned assessment of a situation approaches a high probability of a major harm (category 1) to a third party, the weight of the obligation to breach confidentiality increases. As the situation approaches category 4, breaching confidentiality will likely not be justified. Many particularities of the case will determine whether the professional is justified in breaching confidentiality in categories 2 and 3. These particularities include the foreseeability of a harm's occurrence, the preventability of that harm through a health professional's intervention, the harm that will be caused to the patient, and the potential impact of disclosure on laws and policies regarding confidentiality. However, these abstract conditions are difficult to operationalize, and measurements of probability and magnitude of harm are often not precise.

As a result of these difficulties, efforts have been made to limit the scope of an obligation that requires a breach of confidentiality. Often substituted is a permissive approach that *allows* the clinician to breach confidentiality under some circumstances but does not *obligate* him or her to do so. However, a permissive approach still faces issues about foreseeability and prediction of harms. Enacting statutes to require physicians, psychotherapists, and others to protect third parties from a patient's or client's possible violence, have often limited a Tarasoff-like protective duty, including the duty to warn, to patient-identified or reasonably identifiable third parties. However, the supreme court in the state of Washington has broadened the protective duty beyond specific threats by including more indefinite, ill-defined, and vague threats.[74]

Notification of partners of HIV-positive persons. We will now consider how to specify and balance protections of confidentiality with the protection of third parties in deliberations about whether to disclose a patient's HIV status or genetic information to third parties who might use the information to avoid or mitigate harm.

Physicians and medical institutions must report cases of HIV-infection to public health departments in most legal jurisdictions in the United States. Notably controversial since the beginning of the HIV/AIDS epidemic in the early 1980s is whether physicians and other health care professionals may or should notify at-risk persons that a patient has tested positive for HIV infection and therefore has the potential to infect others. In one case, after several weeks of dry, persistent coughing and night sweats, a bisexual man visited his family physician, who arranged for a test to determine whether he had antibodies to HIV. The physician informed the patient of a positive test, of the risk of infection for his wife, and of the risk that their children might lose both parents. The patient refused to tell his wife and insisted that the physician maintain strict confidentiality. The physician reluctantly agreed. Only in the last few weeks of his life did the patient allow the physician to inform his wife of the nature of her husband's illness, and a test then showed that she too was antibody-positive for HIV. When symptoms appeared a year later, she angrily—but appropriately—accused the physician of violating his moral responsibilities to her and to her children.[75] In this case there was a high probability, under conditions of unprotected sexual intercourse, of a major harm to an identified individual, which is the paradigm case of conditions that justify a breach of confidentiality.

Many well-grounded reasons support notifying spouses and sexual partners—in some cases—that a person has tested positive for HIV. For example, if people are at risk of serious harms, and the disclosure is necessary to prevent and probably would prevent the harms (to their spouses or lovers), then disclosure that breaks confidentiality is warranted. Variations on these conditions appear in several statements of professional ethics by medical associations, but ambiguities and gaps in their statements indicate some difficulties in precisely specifying the nature, scope, and strength of a clinician's moral obligations to protect third parties. Guidelines often do not oblige the physician to determine whether the patient has, in fact, carried out a promise to terminate risky conduct or to warn those endangered, and it is not clear how far the physician or other health care providers should go in monitoring compliance, particularly without the patient's consent.

According to one study, leaving partner notification to patients, sometimes called "passive notification" or "passive referral," is ineffective.[76] However, other studies have found it to be beneficial in getting partners tested and identifying HIV-positive partners, though less effective in linking them with treatment.[77] The

generally preferred and recommended approach, based on several studies, is "assisted partner notification," in which the health care provider assists the patient/client in a disclosure to or anonymous notification of the partner(s) or, with the patient's consent, contacts the partner(s) directly. In one version of "assisted partner notification," called "contract referral," the HIV-positive person establishes a contract with a health care provider to inform his or her sexual or drug-injecting partners directly or anonymously of their probable exposure to HIV. If the partner or partners fail to contact the health care provider within a predetermined period, the health care provider has the right and obligation to make the contact.[78]

Rules of partner notification, whether undertaken by public health officials or other health professionals, changed from their earlier versions in part because (1) HIV/AIDS is now considered a chronic disease rather than a death sentence, (2) there are powerful and effective anti-retroviral therapies that greatly reduce patient mortality, morbidity, and risk to others, as well as other treatments for associated problems, and (3) the stigma of HIV infection and AIDS has declined. HIV partner notification is often incorporated into or operates in close parallel with partner notification for other sexually transmitted infections. Given that as many as 40% of people with HIV infection around the world are undiagnosed, HIV partner notification aims at informing at-risk individuals, inviting them for testing, starting medical therapy for those who test positive for the virus, and educating them as well as those who test negative about practices to reduce the risk of HIV infection. Effective anti-retroviral therapy is also a form of prevention of the spread of HIV infection. Identifying those who have undiagnosed HIV infection is a way to link them to treatment and care and to reduce their risk of infecting others.

The AMA's Council on Ethical and Judicial Affairs has proposed a strategy for dealing with patients with HIV who are putting their partners at risk: Physicians should "attempt to persuade patients who are identified as HIV positive to cease endangering others" and should "be aware of and adhere to state and local guidelines regarding public health reporting and disclosure of HIV status when a patient who is identified as HIV positive poses significant risk of infecting an identifiable third party. The doctor may, if permitted [by state law], notify the endangered third party without revealing the identity of the source person."[79] This strategy limits the obligation to an identifiable third party and permits, but does not require, notification of the third party, without, technically, breaching confidentiality by disclosing "the identity of the source person." Unfortunately, in many cases informing a person that he or she is at risk is sufficient for him or her to identify the "source person."

In contrast, World Health Organization (WHO) guidelines stress that "partner notification services should always be voluntary."[80] A reason for emphasizing voluntariness is that the index patient's cooperation is needed to identify partners who should be notified. The main barriers to effective partner notification services include the index patient who declines to provide information or share partners' names or whose partners were anonymous.[81]

As we have seen, some recommendations and guidelines stress the ethical *permissibility* of the physician's disclosure, whereas others focus on its *obligatoriness*. We need not choose between these two approaches because they can be rendered coherent. There are circumstances in which it is obligatory to so act, centered on our category 1 in the previous chart, whereas other circumstances render it either permissible to so act or permissible not to so act. These are likely cases that fall in categories 2 and 3 that also border on category 1. Assessments of probability and magnitude of harm often do not explicitly indicate whether unconsented-to partner notification is obligatory even when permitted. The justification for disclosure is the same in both sorts of case, namely, reduction of a risk of grave injury or death; but levels of risk and the possibility of a physician's effective action will vary from case to case. (These justificatory conditions are discussed in our requirements of constrained balancing introduced in Chapter 1, pp. 22–24.)

In conclusion, disclosure to an at-risk third party is a morally worrisome act that challenges a central, long-standing professional obligation of confidentiality,[82] though historically this obligation has often not been considered absolute within the medical profession.[83] As a matter of public policy, officials must consider both the critical need to protect endangered third parties and the impact of flexible or rigid societal rules of

confidentiality—including which rules of confidentiality will save more lives in the long run, given that effective protection depends on an individual's provision of information about contacts.

Disclosure of genetic information to third parties. Another ethical problem about notifying at-risk parties arises when physicians, genetic counselors, and others have genetic information about a particular individual that stands to reveal important information about other family members. Individuals who learn that they have a serious genetic condition may have a moral obligation to share that information with at-risk relatives, who may then be able to take actions to reduce risks to themselves or their offspring or to seek treatment. Health care providers should stress this obligation to their patients or clients. Genetic counselors, in particular, may have to overcome their understandable proclivity for nondirective counseling and seek to persuade counselees to disclose relevant information, even though in some ways it would be preferable for the counselors to make the disclosure to ensure that adequate information is transmitted about risks and preventive or therapeutic options.

However, directive counseling differs from disclosure of the information to relatives against the counselee's explicit directive. We concur with the recommendation of the US Institute of Medicine Committee on Assessing Genetic Risks that "confidentiality should be breached and relatives informed about genetic risks only when (1) attempts to elicit voluntary disclosure fail, (2) there is a high probability of irreversible or fatal harm to the relative, (3) the disclosure of the information will [likely] prevent the harm, (4) the disclosure is limited to the information necessary for diagnosis or treatment of the relative, and (5) there is no other reasonable way to avert the harm."[84]

This recommendation closely matches our general approach of constrained balancing. Health care professionals have a prima facie obligation to respect the confidentiality of an individual's personal genetic information, but in certain circumstances they have a right and sometimes an obligation to disclose that information to protect others from harm even if the first party objects. As a general rule, when there is no appropriate consent, the default is nondisclosure to family members at risk.

Some critics of this approach propose that we take more seriously the familial nature of genetic information.[85] By analogy with a bank account, they recommend a model of genetic information as a joint account, whereas we regard it as a personal account. The personal account model fits well with respect for autonomy, confidentiality, maintenance of trust in health care relationships, and good practice in most of health care.[86] In the joint account or relational model, the default is the availability of genetic information to everyone on the account. The default is followed unless there are good reasons not to do so, such as the probability of serious harm to the individual from whom the genetic information was generated. In this approach, disclosure of genetic information to relatives would not necessarily be a breach of confidentiality in a strict sense even if the source individual opposed it, because the genetic information belongs to the relatives too. In some cases, it would be possible and preferable to alert an at-risk relative, for instance, of the need to test for Huntington's disease because of a family history without specific reference to the index case.[87]

A justification of the joint account can be grounded in considerations of justice and reciprocity-based beneficence. Its premise is that one family member should not be able to benefit from jointly valuable information while excluding others from that information and its benefits. Relevant empirical research indicates that many patients view genetic information, such as a genetic mutation, as familial, but consider other aspects of genetic conditions, such as their effect on every-day health, as personal.[88] Patients are usually willing to disclose or to allow disclosure of narrowly shaped genetic information, especially to close relatives.

Nevertheless, if a joint account model, with a default of disclosure, is adopted, it is morally obligatory to inform users of genetic services, at the point of entry, about the nature and limits of confidentiality, so they can choose whether to proceed. The principle of respect for autonomy remains central to any ethically justified use of the joint account model. Rather than changing the default about sharing of information with family members, it would be wiser to further educate individuals about their responsibilities to the family members who could benefit or avoid harm by obtaining this genetic information and then to help facilitate communication of the information.[89]

To reach beyond this example of genetic information: Informing and educating prospective and current patients about the limits of confidentiality of the information generated in health care and about the conditions under which such information may or will be disclosed to others is a vital service, even if patients occasionally object. Given our knowledge that many patients have robust expectations about confidentiality, they should be informed by mental health professionals and others about so-called conditional confidentiality.[90]

FIDELITY

According to Paul Ramsey, the fundamental ethical question in health care and research is, "What is the meaning of the faithfulness of one human being to another?"[91] Few today would agree that fidelity is the fundamental moral norm in health care and research, but it remains a central and often underappreciated moral norm.

The Nature and Place of Fidelity

Obligations of fidelity arise whenever a physician or other health care professional establishes a significant fiduciary relationship with a patient. For our purposes *fiduciary obligation* entails that the professional is morally obligated to act faithfully for another's benefit. To establish a fiduciary relationship is to give an explicit or implicit promise to faithfully carry out or refrain from carrying out an activity. Abandonment of a patient is an example of a breach of fidelity that amounts to disloyalty. Obligations and virtues of fidelity are important at the intersection of research ethics and clinical ethics, where conflicts involving divided loyalties can arise, and we begin with these conflicts.

Conflicts of fidelity and divided loyalties. Professional fidelity, or loyalty, has been traditionally conceived as prioritizing the patient's interests in two respects: (1) the professional effaces self-interest in any conflict with the patient's interests, and (2) the professional favors the patient's interests over third-party interests. In practice, however, fidelity has rarely, if ever, been so pristine. For instance, caring for patients in epidemics has often been considered praiseworthy and virtuous rather than an obligatory enactment of fidelity, and physicians have never been expected to care for a great many patients without compensation. Health care professionals also regularly use their clinical skills to serve social purposes beyond the individual patient's interests, including the protection of public health. They may, for instance, recommend vaccination when, in a context of high rates of immunization, its risks would outweigh its benefits to certain patients. Clinical skills also sometimes serve non-health-related social activities, such as criminal justice and war, as well as religious and cultural practices such as infant male circumcision. Finally, physicians sometimes serve as gatekeepers in society, and in this role divided loyalties are troublesome even when unavoidable. Examples include providing psychiatric evaluation as part of a criminal trial, performing employee reviews for an employer (as in the Bratt case examined earlier), and conducting a medical review of a person's disability insurance claims.[92]

Divided loyalties typically occur when fidelity to patients, subjects, or clients conflicts with allegiance to colleagues, institutions, funding agencies, corporations, or the state. Conflicts in dual roles are intensely felt in fields such as forensic medicine and military medicine. In these conflicts, two or more roles and their coupled loyalties and obligations are sometimes incompatible and irreconcilable, forcing a moral choice.[93]

Third-party interests. Physicians, nurses, and hospital administrators sometimes find aspects of their role obligations in conflict with their obligations to patients. In some cases, they have a therapeutic contract with a party other than the patient. For instance, when parents bring a child to a physician for treatment, the physician's primary responsibility is to serve the child's interests, even though the parents made the contract and the physician has obligations of fidelity to the parents. The latter obligations are sometimes validly overridden, as occurs when physicians go to court to oppose parents' decisions that seriously threaten their children. Courts have allowed adult Jehovah's Witnesses, for example, to reject blood transfusions for themselves as competent adults while disallowing parental rejections of medically necessary blood transfusions for their children. Parents are sometimes appropriately charged legally with child neglect when they fail to seek or permit potentially beneficial medical treatment recommended by physicians.[94]

Institutional interests. In some conflicts, it is unclear what the health care professional owes the "patient." The institutions involved may not be health care institutions, but, in discharging their functions, they may need medical information about individuals and may even provide some care for those individuals. Examples include a physician's contract to provide medical examinations of applicants for positions in a company or to determine whether applicants for insurance policies are safe risks. In some circumstances the health care professional may rightly not regard the person examined as his or her patient, but, even so, the professional has moral responsibilities of due care including disclosure of serious risks ascertained through the medical examination.

In some jurisdictions the health care professional does not have a legal obligation to disclose the discovery of a risk or of a disease to the examinee, but nondisclosure may be a morally dubious practice. At a minimum, health care professionals have a moral responsibility to oppose, avoid, and withdraw from contracts that would require them to withhold vital health information from examinees. Physicians often have "due care" obligations to individuals who become their patients under a third-party contract in an institutional arrangement. Examples include industries, prisons, the armed services, and professional sports teams.

Nevertheless, when care of an individual conflicts with institutional objectives and policies to which a health professional is also committed, the individual's needs do not always take precedence. For example, a military physician must accept a different set of obligations than those of a nonmilitary physician to place the military's interests above both the patient's and the physician's interests. A moral dilemma may emerge for the military physician in determining whether to certify a soldier suffering from a closed-head injury, resulting from an improvised explosive device, as fit to return to the front lines. On the one hand, the soldier, while medically stable and functional, continues to experience fatigue, problems in sleeping, and daily headaches and would be at increased risk of worse impairments and post-traumatic stress syndrome if a similar incident occurred again. On the other hand, his commanding officers have indicated their critical need for his particular expertise and experience.[95] Apart from such dilemmas, some actions so grossly violate canons of medical ethics that they warrant disobedience of orders and defiance of superiors rather than loyalty and compliance. An example is a commander's order for a physician to help torture a prisoner of war.[96]

Medical assistance in prisons also presents moral challenges, in part because of the institutional mandate to punish the criminal, which limits the obligations of fidelity to the criminal as patient. Medical values are sometimes subordinated to the correctional institution's functions, but the physician is expected to be loyal to both. The correctional institution may expect physicians and other health care professionals to participate in the administration of justice and punishment. Examples include surgical removal of a bullet for evidence when the bullet is not a hazard to the inmate and can be safely left in place, forced examinations of inmates' body cavities for evidence of contraband drugs, and participation in corporal or capital punishment—for instance, by administering a lethal injection.[97] Moral questions also arise both about medical assessments of prisoners' physical conditions in order to determine whether they can endure punishments and about medical monitoring of prisoners during punishment. Such medical assessments and supervision can reduce the likelihood of extreme or unintended injury, but participation in the actual administration of punishment, whether corporal or capital, is a compromise of fidelity.[98]

Nursing. Nursing may be the area of health care with the most pervasive range of conflicts. In the latter part of the twentieth century, codes of nursing ethics began to frame nurses' moral responsibility in ways different from earlier codes, which functioned to discourage nurses from making their own moral judgments. In 1950, the first code of the American Nurses Association stressed the nurse's obligation to carry out the physician's orders, whereas the 1976 revision stressed the nurse's obligations to the patient. The current version emphasizes that the nurse's "primary commitment is to the patient, whether an individual, family, group, community, or population" and that the nurse "promotes, advocates for, and protects the rights, health, and safety of the patient."[99] (See our discussion in Chapter 2, pp. 34–35, of virtues associated with different conceptions of nursing.)

Moral conflicts can be expected wherever one group of professionals makes the decisions and orders their implementation by other professionals who have not participated in the decision making. In one study of relationships in health care, investigators examined different perceptions of ethical problems by nurses and

doctors in acute care units. In structured interviews, both nurses and physicians said they frequently encountered ethical problems. Most physicians (twenty-one of twenty-four) and most nurses (twenty-five of twenty-six) recognized ethical conflicts within the health care team. In twenty-one of the twenty-five cases reported by nurses, the ethical conflict was between a nurse and a physician, whereas only one physician reported a conflict with a nurse rather than with another physician. The authors of the study conclude that it is likely that conflicts with nurses occurred, but that the physicians "were not aware of them, or did not see conflict with a nurse as forming an *ethical* problem."[100] Several features of the working relationship between physicians and nurses help to explain these findings. Physicians commonly write orders and nurses execute them. Through their close relationships with patients, nurses often experience, more immediately than physicians, problems that arise from medical decisions.

According to another study of physicians and registered nurses caring for dying patients in intensive care units (ICUs), the nurses experienced greater moral distress, perceived the moral climate in their ICU more negatively, expressed less satisfaction with the quality of care, and experienced less collaboration than the physicians reported. As a solution, the researchers appropriately proposed not only improving collaboration but also paying explicit attention to situations that generate moral distress and to differences in role perspectives.[101] Other studies find a need to recognize and engage legitimate moral disagreement among different health care professionals, including physicians, nurses, and administrators within a respectful moral community, before moral conflicts arise that may compromise patient care.[102]

Problems of Conflict of Interest

In recent years, traditional rules of fidelity have often been threatened or weakened by conflicts of interest, a fairly recent concern in medicine (and in biomedical ethics) in contrast to other professions such as law. A conflict of interest exists when an impartial observer would determine that a professional's judgments, decisions, or actions are at risk of being unduly influenced by his or her personal interests, such as financial interests or friendship.[103] The risk is that the professional's personal interests will create temptations, biases, and the like that will lead to a betrayal of role responsibilities through judgments, decisions, and actions other than those reasonably expected of a person in the role. The reasonable expectation is that clinicians will seek the patient's welfare and respect his or her rights, that researchers will pursue objective and valid results, and so forth. A conflict of interest creates a risk that the professional in question will compromise these expectations and thereby damage patients' interests and rights, distort research, or teach trainees in a biased way.

The analysis and assessment of risks of different types of conflict of interest follow the risk chart introduced earlier. The degree or level of risk depends on (1) the probability that the professional's personal interests will have an undue influence on his or her judgments, decisions, or actions, and (2) the magnitude of harm that may occur as a result. Even if the circumstance of conflict does not in fact bias the individual's judgment, and even if no wrong is committed, it is still a conflict-of-interest situation that makes it reasonable to suppose that tainted judgments might occur and to require that they be disclosed, mitigated, managed, or avoided altogether.

Conflicts of interest occur in medicine, health care, biomedical research, the development of clinical practice standards, and the review of grant proposals and articles submitted for publication in all of these fields. Although the medical profession has not paid adequate attention to nonfinancial conflicts, such as professional advancement or friendship, which are no less important, numerous efforts are under way to address various financial conflicts, including fee splitting, self-referring, accepting gifts, accepting fees for recruiting patients for a research protocol, outside consulting with a regulated industry by government-employed physicians, appointing industry-based physicians to government regulatory agencies, and industry-paid lecturing on an industry product.

One issue is the referral of patients to medical facilities or services physicians own or in which they have a financial investment. Self-referral threatens fidelity to patients' interests by enlarging the temptation inherent in fee-for-service to provide pointless, optional, or excessively expensive care. Physicians create these financial conflicts of interest by owning or investing in medical facilities or services, such as diagnostic imaging centers,

laboratories, or physical therapy services, to which they then refer patients. Physician ownership of radiation therapy and physical therapy services, for instance, can increase use and costs, without compensatory benefits to patients such as increased access.[104] Self-referral is usually more problematic than fee-for-service because the patient typically cannot identify the physician's potential economic gain—unless it is explicitly disclosed—in ordering additional procedures and thus cannot proceed cautiously, perhaps by seeking a second opinion.

In our judgment, physicians have an ethical obligation to disclose both financial and nonfinancial conflicts of interest, but disclosures are not as widespread as they should be. According to a 2009 national survey, nearly two-fifths of the responding physicians, across several medical specialties, did not entirely agree that they should disclose to patients their financial relationships with drug or device companies.[105] Fidelity and honesty require such disclosure as an ethical minimum, though disclosure is rarely sufficient. For instance, it is unclear how a vulnerable patient can effectively use disclosed information in the context of physician self-referral. Moreover, legal or professional prohibition of self-referral is warranted in many types of cases.[106]

Third-party payers and institutional providers have imposed many constraints on medical decisions about diagnostic and therapeutic procedures through mechanisms designed to control costs. These mechanisms sometimes limit and constrict the physician's ability to maintain fidelity to the patient because of incentives and disincentives that can place the physician's self-interest in conflict with the patient's best medical interest. For example, health maintenance organizations (HMOs) may withhold a substantial part of the primary physician's income. At the end of the year, they return part or all, depending on the overall financial condition of the HMO and, in some cases, the physician's productivity and frugality. This arrangement creates an incentive for physicians to severely limit expensive procedures—a worrisome conflict of interest. The patient is in a markedly different position when the physician has a conflict because of incentives to *limit needed treatment* than when the physician has a conflict because of incentives to *provide unnecessary treatment*. In the latter situation, patients often can obtain another opinion. In the former situation, patients may not be aware of a needed treatment that is omitted.[107] Both are ethically unacceptable, at least when the incentives are likely to influence treatment decisions, and both require corrective measures.

Financial incentive structures such as those used in many diagnostic laboratories also create a motive for physicians to limit both their time and expensive procedures. Physicians are paid by measurable output, and annual payments are tied to productivity measures such as the number of slides read. But a rapid reading of data compounds the risk of error, increasing risks of false-negative results and misdiagnosis. Pathologists who read hundreds of slides per day looking for the presence of carcinoma will boost their salary in these arrangements but also will increase the likelihood of failing to detect a carcinoma. Every physician will occasionally make a mistake or follow an incorrect, yet excusable, strategy; but it is not morally excusable to make mistakes where there is an inherent conflict of interest encouraging behavior that falls below an appropriate standard of due care.

Another set of conflicts of interest arises from gifts from pharmaceutical and medical device manufacturers. In contrast to a widespread assumption that only large gifts create conflicts of interest, there is evidence that even small gifts, such as pens, note pads, and lunches, intended to build and maintain relationships influence physicians' prescribing behavior.[108] Moreover, gift relationships, however small, create a variety of temptations, dependencies, friendships, and forms of indebtedness—all of which stand to create conflicts of interest with the physician's primary obligation to act in the best interest of the patients.[109] Although disclosure to patients may help to reduce the negative impact of several forms of conflict of interest, it appears to be relatively useless for conflicts of interest created by industry gifts to physicians. More stringent regulations by institutions, including academic medical centers, are needed to eliminate, curtail, or modify these common practices in the interactions between, for example, industries and physicians. Institutional rules could include banning gifts, not accepting funds for lunches at educational programs, and reducing the practice of accepting free samples.[110]

Conflicts of interest reach beyond practice to research. Interactions and partnerships involving industry, government, and the academy are vitally important in the development, support, and conduct of biomedical research to benefit human health, and yet they frequently harbor conflicts of interest.[111] For instance, clinical trials of pharmaceutical products are often funded by companies willing to assume the financial risk because the

returns from successful trials are the lifeblood of the company. The joint financial advantages for physician-investigators and corporations promote a relationship that may ensure a steady and reliable funding stream. But this relationship also risks creating a motive for physician-investigators to find positive results or downplay negative results, thereby compromising scientific objectivity. It is vital to control the process of interpretation and assessment through objective procedures, backup checks, and independent controls such as data safety and monitoring boards.[112]

Journals too should require researchers to provide information about sources of funds for the research. Furthermore, as an Institute of Medicine (later the National Academy of Medicine) report recommended, researchers should not conduct research involving human subjects if they have a significant financial stake in the outcome of that research. For example, a researcher might hold a patent on a product being tested in a clinical trial. However, there are good reasons to make exceptions in rare circumstances such as when an institutional conflict-of-interest committee determines that an individual's participation is essential for the safety or validity of the research (perhaps because of the complexity of the procedure or device that the researcher developed) and that it is possible to manage the conflict and ensure the integrity of the research.[113]

Deliberative assessments need to be made regarding ways to address these various types of conflict of interest. For example, we might eliminate them, manage or mitigate them, or require disclosure of conflicts to parties at risk. Each strategy is justifiable in some contexts, and each is preferable to the traditional practice of relying on professional judgment or personal character to determine whether a conflict is *actual*, *potential*, or *merely apparent*—a dubious set of distinctions since the potential, or so-called apparent, conflict often constitutes an actual conflict of interest. Professionals sometimes view attempts to address conflicts of interest as negative judgments on their and their colleagues' character, as though they might be corrupt and might act against the reasonable expectations of their professional roles in the pursuit of personal self-interest, but this assessment misses the point of conflict-of-interest rules. Unconscious and unintentional distortions of professional judgments, decisions, and actions are at least as important. It is difficult to perform individualized assessments of the likelihood that a particular professional's conflicts of interest will lead to a breach of professional expectations. Accordingly, general rules, regulations, and impartial oversight are essential.

THE DISTINCTION BETWEEN CLINICAL ETHICS AND RESEARCH ETHICS

Now that we have completed our examination of rules and virtues of veracity, privacy, confidentiality, and fidelity, we turn to other dimensions of relationships between professionals and their patients or subjects, starting with a basic distinction between clinical research and clinical medicine and how this distinction affects our thinking about professional biomedical ethics.

Biomedical ethics has long drawn a line between clinical (medical and nursing) ethics and research ethics. This line rests on a distinction between clinical practice and clinical research, a distinction that still deeply influences how we conceptualize areas of medicine and biomedical science and how we understand the ethical rules appropriate to them. The research-practice distinction also affects how we think about activities that are subject to governmental regulation. Research has been heavily regulated in many countries because it has been thought to place subjects at risk for the benefit of others and to investigate unconfirmed hypotheses about diagnoses and treatments. By contrast, medical practice is far less regulated on grounds that it focuses on the patient's best interests and relies on interventions of proven benefit and acceptable risk.

This distinction also determines which activities *must* undergo review by an ethics committee. The general conception has long been that if a component of research introduces risk in an activity involving a human person, the component must undergo review to protect subjects. Nothing comparable exists at the national level in most countries for medical practice. But the question is whether such a sharp distinction between research and practice, as well as the parallel differences in ethical and regulatory frameworks, are truly warranted. Why, morally, should practice be treated so differently from research when it comes to the oversight and protection of clinical patients and research subjects?

The Place and Significance of "Generalizable Knowledge"

The conventional assessment has been that research lacks a focus on personalized care. Its distinctive objective is scientifically designed testing of a hypothesis aimed at developing or contributing to what US regulations—and the bioethics literature generally—refer to as "generalizable knowledge."[114] By contrast, medical practice interventions are aimed at diagnosis, preventive treatment, or therapy that stands to provide optimal therapeutic benefit to each patient. In clinical medicine, risks are justified by an intervention's potential benefit to the individual patient, whereas in clinical research, risks are usually justified by the potential social benefit of the research (sometimes combined with a possible benefit to the patient). In this conception, risk allocation sharply differs in the clinic and in research. This idea has supported the view that clinical research and clinical practice require different ethical rules, in accordance with the different objectives, roles, and relationships that characterize each. Accordingly, there are distinct particular moralities—clinical ethics and research ethics—each with its own system of moral norms and review.[115]

This entrenched distinction between research and practice is puzzling and morally questionable. First, the boundaries between research and practice are often porous, especially when the two occur together in the same health care institution(s), and each contributes to the other. A good example is pediatric oncology and its remarkable practical successes in treatment in recent decades. In this field of medicine, research is intimately intertwined with practice, and practice is often not divorced from new infusions of data from research. Second, some areas of medicine use innovative techniques or practices that have never been scientifically validated through research and lack regulatory approval for these uses. These practices are often regarded, rightly, as experimental, which suggests that patients so treated *are* subjects of "research," even if no effort is made to produce generalizable knowledge, but only an effort to benefit the patient.[116]

Innovative treatments, including off-label uses of treatments (e.g., uses of prescription drugs to treat conditions for which the drug has not been officially approved), fall short of the high validation standards set by randomized clinical trials. Although the range of acceptable methods of obtaining knowledge in medicine is controversial, it is morally unsatisfactory to allow physicians to use treatments that are either unapproved or new on grounds that the patient-physician relationship is a private transaction immune from regulatory interference and unaccountable to external oversight, such as a review committee. Many parts of medical practice conform to this model, but, in general, there is no reason to think that well-designed clinical research is riskier than standard forms of clinical practice based on innovative therapies.

Is Our Ethical Oversight of Research and Practice Coherent?

Accordingly, we need to examine whether we have a coherent moral conception of the ethical oversight of research and practice. The central ethical issue is whether research projects require the higher level of scrutiny that now occurs and whether clinical encounters need a lower level of scrutiny. If the risks are similar and the need for consent to interventions is similar, arguably the oversight system should be relevantly similar, regardless of the conventional categorizations of research and clinical practice. The time is now ripe for a closer and more thorough examination of these categories and distinctions in biomedical ethics and public policy.[117]

Whatever the appropriate level of scrutiny and oversight of each of these activities, the dual roles of clinician and investigator generate possible conflicts of obligation and of interest that require attention, as we will now argue.

The Dual Roles of Clinician and Investigator

The "Declaration of Geneva, Physician's Pledge" of the World Medical Association, as revised in 2017, affirms that "the health and well-being of my patient will be my first consideration."[118] But can research involving patients and other subjects or participants consistently honor this obligation? The dual roles of research scientist and clinical practitioner pull in different directions, potentially creating significant conflicts of obligation and of

interest. As an investigator, the physician acts to generate scientific knowledge to benefit individual patients and populations, usually in the future. As a clinician, the physician has the responsibility to act in the best interests of present patients. Accordingly, responsibilities to future generations may conflict with due care for current patients who become research subjects.

Research involving human subjects is a vital social enterprise to advance science, but to be ethically justified it must satisfy several conditions, including (1) a goal of valuable knowledge, (2) a reasonable prospect that the research will generate the knowledge that is sought, (3) the necessity of using human subjects, (4) a favorable balance of potential benefits over risks to the subjects, (5) fair selection of subjects, and (6) measures to protect privacy and confidentiality. Only if these conditions have been met is it appropriate to invite potential subjects (or their surrogates) to give their informed consent or refusal to participate. Consent can therefore be considered a seventh condition that must be met.[119]

These conditions apply to both research that offers no prospect of direct medical benefit to the subject and research that offers some prospect of direct medical benefit to the patient-subject and that may be conducted during the course of the care of the patient. The term *therapeutic research* is potentially misleading because, when misunderstood, it can draw attention away from the fact that research is being conducted. Clinical research is distinguishable from both routine therapy and experimental or innovative therapy, which are directed at particular patients. Attaching the term *therapeutic* to research may create a "therapeutic misconception," in which participants construe the protocol design as therapy directed at the individual rather than as research designed to generate generalizable knowledge. (See further the discussion of therapeutic misconception in Chapter 4, pp. 132–33.)

Because society encourages and supports extensive research and because investigators and subjects are unequal in knowledge and vulnerability, public policy and review committees are responsible for ensuring that the research meets the several conditions noted above. Some cases warrant a straightforward paternalistic decision. For example, if healthy persons free of heart disease volunteer to participate in a research protocol to test an artificial heart, as once happened,[120] an institutional review board (IRB) should declare that the risk relative to benefit for a healthy subject is too substantial to permit the research to enroll such subjects, whereas the risk relative to benefit for a patient with a seriously diseased heart may be acceptable.

Conflicts in Clinical Trials

Controlled clinical trials are often essential to establish or confirm that an observed effect, such as reduced mortality from a disease, results from a particular intervention rather than from an unknown variable in the patient population. The evidence supporting many available treatments is tenuous, and some may have never been adequately tested for either safety or efficacy. Even if adequate testing occurred at one time, the treatments may no longer be as safe or as efficacious as new treatments—a matter of their comparative effectiveness. If doubt surrounds the efficacy or safety of a treatment, or its relative merits in comparison to another treatment, scientific research to resolve the doubt is in order.[121]

Controlled trials are scientific instruments intended to protect current and future patients against medical enthusiasm, hunches, and outdated procedures and products. In these trials, one group receives the investigational (or experimental) therapy, while a "control group" receives either a standard therapy or a placebo (an inert preparation that resembles a drug in appearance) so that investigators can determine whether an investigational therapy is more effective and safer than a standard therapy, placebo, or no treatment. Commonly, subjects are randomly assigned to either control or investigational groups to avoid intentional or unintentional bias. Randomization is designed to keep variables other than the treatments under examination from distorting study results.

Blinding certain persons to some information about the randomized controlled trial (RCT) provides additional protection against bias. An RCT may be single-blind (the subject does not know whether he or she is in the control group or the experimental group), double-blind (neither the subject nor the investigator knows), or unblinded (all parties know). Double-blind studies are designed to reduce bias in observations and

interpretations by subjects, physicians, and investigators. Blinding the physician-investigator also serves an ethical function, because it partially obviates the conflicts of obligation and of interest that arise for physicians who are simultaneously engaged in clinical practice and research with the same patient(s).

Problems of consent. By design, subjects in RCTs usually do not know which treatment or placebo they will receive. However, no justification exists for not disclosing to potential subjects the full set of methods, treatments, and placebos (if any) that will be used, their known risks and probable benefits, and any known uncertainties. Likewise, no justification exists for failing to disclose the rationale for the study, the fact of randomization, how the trial differs from clinical practice, and available alternatives to participation. A physician-researcher with dual responsibilities also has a fiduciary obligation to inform patient-subjects of any relevant conflicts of interest.[122] With this information, potential subjects will almost always have an adequate basis for deciding whether to participate.

In conventional RCTs, investigators screen patients for eligibility and then provide the information just discussed. If a patient consents to participate, he or she is then randomized to one arm of the study. However, even when scientific evidence indicates that two proposed interventions are roughly equal in safety and efficacy, patients may have a strong preference for one arm over another. Consider a situation in which two surgical procedures for treating the same disease appear to have the same survival rate (say, an average of fifteen years) and we want to test their effectiveness by an RCT. A patient might have a preference if treatment A has little risk of death during the operation but a high rate of death after ten years, and treatment B has a high risk of death during the operation or post-operative recovery but a low rate of death after recovery for thirty years. Factors such as a patient's age, family responsibilities, and other circumstances might lead to a preference for one over the other. Accordingly, some patients may choose not to enter a particular RCT even though, from the standpoint of safety and efficacy, the different arms are in clinical equipoise—our next topic of discussion.

The problem of clinical equipoise. Serving the patient's best interests is intuitively inconsistent with assigning a treatment randomly in order to promote social goals of accumulating knowledge and benefiting future patients. It seems inconceivable that optimal medical care occurs by random assignment to an intervention or no intervention. No two patients are alike, and ordinarily a physician should be able to select and modify the course of therapy, as needed, to promote the patient's best interests. The question is whether this traditional axiom of medical ethics is consistent with RCTs.

Proponents argue that RCTs do not violate moral obligations to patients because they are used only in circumstances in which justifiable doubt exists about the relative merits of existing, standard, and new therapies. No one knows, prior to conducting the research, whether it is more advantageous to be in the control group or in the experimental group. The community of reasonable physicians is therefore in a state of "clinical equipoise."[123] On the basis of the available evidence, members of the relevant expert medical community are uncertain or disagree about which intervention is superior and so are equally poised between the treatment strategies under examination in the RCT. That is, they are equally uncertain or disagree about the known advantages and disadvantages of the investigational treatment to be tested and the current treatment, placebo, or no treatment the control group will receive. In this model, no patient will receive something known to be less effective or to have a higher risk than an available alternative.

When patients are not asked to forgo a superior treatment, the use of RCTs is justifiable, especially in light of the promise of benefit to future patients. In the absence of scientific grounds before the trial for preferring to be in one group rather than another, a patient may prefer one over the other on the basis of hunches or intuitions about effectiveness and safety or on the basis of factors not being studied in the trial. If two treatments for breast cancer, for example, are in veritable clinical equipoise from the standpoint of survival, a woman still may prefer the less disfiguring treatment.

Some critics of appeals to clinical equipoise as a way to establish the moral legitimacy of clinical trials are concerned about an excessively narrow focus on the ethics of the clinical physician's role, chiefly on whether RCTs are consistent with physician duties in the physician-patient relationship. This approach, some critics maintain, neglects society's considerable interest in evidence-based health policy and advances in scientific

understanding needed for drug approval and coverage decisions.[124] This concern about the general justification of RCTs offers a fair warning about the need to avoid an unduly narrow focus, but it does not negate the need to investigate ethical conflicts in clinical trials, which is our focus in this section. Clinical equipoise is an important threshold condition that must be met in the conduct of RCTs. However, it is not a sufficient condition of the moral legitimacy of RCTs and does not by itself adequately guide social policy. Whether particular RCTs in fact satisfy this threshold condition is understandably debated, as is appropriate governance of the research through social policies.[125]

Finally, if a cooperating physician strongly believes prior to a trial that, based on available evidence, one therapy is more beneficial or safer, he or she must decide whether to suspend this belief in the interests of scientific objectivity and in deference to the views of the community of experts, who find themselves in clinical equipoise. In this circumstance, as part of the informed consent process, the physician is morally obligated to disclose both his or her personal conviction and that of the relevant community of experts to patients who are potential candidates for the trial.[126]

The problem of placebo controls. Conducting placebo-controlled and no-treatment trials is controversial, especially when an established and effective treatment exists for the condition under investigation. Critics argue that the use of placebo controls is unethical because placebo-controlled trials may not be methodologically superior and may deny patients treatment when treatment controls (a group that receives an established effective intervention, called *active controls* in the literature) could be used.[127] By contrast, defenders of placebo-controlled trials contend that they are methodologically superior to active controlled trials and are frequently essential in the process of scientific validation.[128] Fortunately, all parties agree that use of a placebo is ethically acceptable only if there is a reasonable prospect of producing scientifically valid information by this method. As best we can determine, placebo-controlled trials are often methodologically superior to, more efficient than, and less costly than active controlled trials. They can even be necessary to distinguish treatment effects. Nonetheless, there are moral problems with their use.

The best strategy in addressing this controversy is to locate the conditions under which placebo use is ethically acceptable and the conditions under which it is unacceptable. As a start, if an established effective intervention is available for use in the population to be studied, using a placebo control is unethical if withholding the effective intervention from subjects has a significant probability of being life-threatening, of causing permanent damage, of causing irreversible disease progression, or of causing an unacceptable level of pain or suffering. Placebo use is impermissible under these conditions because the risks of the research to subjects are too high, exceeding a threshold, or the overall benefits do not outweigh the risks to subjects.

By contrast, if no established safe and effective intervention exists to treat the medical problem under study, the use of a placebo is permissible in research on a new investigational therapy. In some cases, the relevant expert community may have significant doubts about the benefits provided by approved available treatments or many patients may not be able to use the available treatment(s) because of their medical condition. Ethical acceptability may also be contingent on other conditions. For example, patients may have refused an established effective treatment under conditions in which withholding that treatment will not cause serious or irreversible harm. Here again use of a placebo may be justified.

In a now classic case of a questionable use of placebo, a conflict erupted over placebo-controlled trials of AZT (azidothymidine) in the treatment of AIDS. Promising laboratory tests led to a trial (phase I) to determine the safety of AZT among patients with AIDS. Several patients showed clinical improvement. Because AIDS was then considered invariably fatal, many people argued that compassion dictated making it immediately available to all patients with AIDS and, perhaps, to those who were antibody-positive to the AIDS virus. However, the pharmaceutical company (Burroughs Wellcome Company, later GlaxoSmithKline Pharmaceuticals) did not have an adequate supply of the drug to satisfy this plan, and, as required by federal regulations, it used a placebo-controlled trial of AZT to determine its effectiveness for certain groups of patients with AIDS. A computer randomly assigned some patients to AZT and others to a placebo. For several months, no major differences emerged in effectiveness, but then patients receiving the placebo began to die at a significantly higher rate. Of

the 137 patients on the placebo, 16 died. Of the 145 patients on AZT, only 1 died.[129] Many moral problems surround starting such a placebo-controlled trial when a disease appears to be universally fatal and no promising alternative to the new treatment exists, and related questions arise about when to stop a trial as well as how to distribute a new treatment.

A second example comes from RCTs in surgery, which are rare, particularly when placebos are used. There are concerns that surgical procedures are too easily introduced without sufficiently rigorous evidence of their efficacy or safety. In one case, surgical researchers sought a clinical trial to determine whether transplanting fetal neural tissue into the brains of patients with Parkinson's disease (a disorder of motor function, marked by tremor, rigidity, unsteady walking, and unstable posture) would be safe and effective. Standard medical treatment consisted of levodopa, which might not restore lost motor function, might have adverse effects over a long period, and might not adequately control new manifestations of the disease. Researchers argued that surgical therapy using cells is more like the administration of pharmaceutical agents than like conventional surgical procedures. In proposing a randomized, double-blind, placebo-controlled trial, they maintained that a placebo control was scientifically preferable to the use of standard medical treatment as the control because surgery itself may have some effects, such as evoking patients' favorable subjective responses. The placebo consisted of sham surgery, that is, the administration of general anesthesia followed by bilateral surgery, which involves a skin incision with a partial burr hole that does not penetrate the skull's inner cortex. This sham surgery was to be compared to two other procedures that differed from each other only in the amount of fetal tissue transplanted. All participants received antibiotics and immunosuppression medication for six months. The thirty-six subjects in this study all knew that twelve of them would undergo sham surgery and researchers promised all of them free access to the real surgery if the trial demonstrated its net benefits.[130]

The main argument against the use of sham surgery as a placebo control in this research focuses on the substantial risks from the procedure and the anesthesia. The best research design, from the standpoints of both the investigators involved and future patients, conflicted with investigators' obligations of beneficence and nonmaleficence to current patients invited to serve as research subjects. Among several ethical questions, one is whether the patient-subjects' informed consent[131] was sufficient to justify proceeding with the research.[132] It is doubtful that informed consent is sufficient by itself in such cases. Consent should be considered together with the level of risk involved, the need to reduce bias by blinding participants, the alternatives that might obviate the need for sham surgery, and the like.

Nonetheless, if we assume that other conditions for ethically justified research are met, genuinely informed consents go a long way toward justifying the conduct of placebo-controlled trials when prospective subjects are informed about the following: A placebo will be used, subjects could be randomized to a placebo arm, the reasons that using a placebo is part of the design, the benefits and risks of already available treatments, the risks of refusing those treatments, the option of receiving the treatment if symptoms worsen, and the right to withdraw for any reason and at any time from the study. Understanding these elements of the research is a necessary condition of an adequately informed consent in this context, but even an elevated informed consent does not by itself justify use of placebo-controlled trials.

Early Termination of and Withdrawal from Clinical Trials

Physician-researchers sometimes face difficult questions about whether to stop a clinical trial before its planned end—particularly whether to withdraw patient-subjects from the trial before sufficient scientific data are available to support definitive conclusions. Access to data is limited during clinical trials to protect the integrity of the research. Consequently, physicians may be excluded from access to critical information about trends. If they became aware of trends prior to the point of statistical significance, they might pull their patients from the trial, and numerous withdrawals might invalidate the research.

However, if a physician determines that a particular patient's condition is deteriorating and that this patient's best interests are served by withdrawal from the research, the physician morally must be free to act on behalf of the patient and recommend withdrawal. In an RCT, it may be agonizingly difficult to determine whether the

research as a whole should be stopped, even if some physician-researchers are satisfied by what they have observed. One procedural solution is to differentiate roles, distinguishing between the responsibilities of individual physicians who must make decisions regarding their own patients and those of a data and safety monitoring board (DSMB) established to determine whether to continue or stop a trial. Unlike patients' physicians, the DSMB is charged to consider the impact of its decision on future patients as well as on current patient-subjects. One of its functions is to stop or recommend stopping a trial if accumulated scientific data indicate that uncertainty has been reduced and equipoise no longer prevails,[133] as happened in the original AZT trial for AIDS. In order to ensure the integrity of the clinical trial, the DSMB must be independent of the investigators and sponsors and able to make objective, impartial analyses, judgments, and recommendations.[134]

This differentiation of roles by using a DSMB is procedurally sound, but it relocates, rather than resolves, some ethical questions. The DSMB must determine whether it is legitimate to impose or to continue to impose risks on current patient-subjects in order to establish a higher degree of probability of the superiority of one treatment over another. It should determine whether clinical equipoise has been disturbed (i.e., eradicated) from the perspective of impartial observers in the expert medical community.[135] However, the individual physician and his or her patient will be primarily concerned with whether clinical uncertainty (and equipoise) has been eliminated or substantially reduced *for them*.

Many questions are relevant to a patient-subject's decision to withdraw from an RCT based on such information, including questions about interim data and early trends. Trends are often misleading and sometimes prove to be temporary aberrations, but they might be relevant at a given point to a patient-subject's decision about whether to continue to participate even if the evidence does not satisfy statisticians or the expert medical community. If information about trends is not to be released prior to the completion or early termination of the RCT, potential subjects need to be informed of this rule and accept it in the informed consent process as a condition of participation.

Justifying Conditions for Randomized Clinical Trials

Despite the several problems we have identified, RCTs are justified—including those involving placebo controls—if they satisfy the following seven substantive and procedural conditions (in addition to the general conditions of justified research previously identified):[136]

1. 1. Clinical equipoise genuinely exists in the community of relevant and impartial medical experts.
2. 2. The trial is designed as a crucial experiment to determine whether an investigational therapeutic alternative is superior to available alternatives and shows scientific promise of achieving this result.
3. 3. An IRB or its functional equivalent has approved the protocol and certified that no physician-investigator has a conflict of interest or incentive that would threaten either the patient-physician relationship or impartiality in the conduct of research.
4. 4. Patient-subjects have given a genuinely informed consent (as we analyze this concept in Chapter 4).
5. 5. Placebos and no-treatment options cannot be used if an effective treatment exists for the condition being studied and that condition threatens death, grave injury, or serious morbidity in the absence of the treatment.
6. 6. A data and safety monitoring board either will end the trial if statistically significant data disrupt clinical equipoise or will supply physicians and patients with substantive safety and therapeutic information that has emerged and shown to be relevant to a reasonable person's decision to remain in or to withdraw from the trial.
7. 7. Physicians have the right to recommend withdrawal and patients have the right to withdraw at any time.

CONCLUSION

In this chapter we have interpreted and specified the principles and virtues of respect for autonomy, nonmaleficence, beneficence, and justice, as analyzed in the previous four chapters. We have concentrated on

obligations and virtues of veracity, privacy, confidentiality, and fidelity, and we have explored the basis, meaning, limits, and stringency of the obligations that govern contexts of professional–patient and professional–subject relationships—and in some cases professional–professional relationships, such as those between physicians and nurses. Similarly, we have shown that moral virtues are often as important as moral obligations for health professionals grappling with these moral problems. These virtues include care, compassion, and discernment as well as the specific virtues examined in this chapter of truthfulness, respect for privacy, confidentiality, and fidelity, each of which we linked to moral obligations (a correspondence discussed in Chapter 8, pp. 414–15).

We have now completed our discussion in Part II of this volume of the four clusters of principles of biomedical ethics, their corresponding virtues, the rules derivative from these principles, and their implications for professional ethics. In the final two chapters, in Part III, we will examine the place of ethical theory and method in biomedical ethics.

NOTES

1. 1. We use the term *patient* in the title of this chapter although, in many of the relationships we discuss, *patient* is not the most accurate term. We make clarifications and qualifications as the chapter proceeds. For human beings who are being studied in research we use the label *human subjects*, in line with tradition and federal regulations. However, in recent bioethical discourse, the term *participant* has become widely used in order to highlight voluntary partnership and cooperation in research. No term is perfect. To many, "subject" suggests being subjected to or under the control of others, while "participant" neglects that many are enrolled in research by others, such as parents, rather than being voluntary participants. See the discussion of this problem in National Bioethics Advisory Commission (NBAC), *Ethical and Policy Issues in Research Involving Human Participants*, vol. 1: *Report and Recommendations* (Bethesda, MD: NBAC, August, 2001), chap. 1, fn 1.
2. 2. *Code of Medical Ethics of the American Medical Association*, 2016–2017 Edition (Chicago: AMA, 2017), p. 1—the Principles of Medical Ethics being "the primary component of the Code"; and *Current Opinions of the Judicial Council of the American Medical Association* (Chicago: AMA, 1981), p. ix. For the original 1847 code, see American Medical Association, *Code of Medical Ethics* (Chicago: AMA, 1847), p. 88, available at https://www.bioethicscourse.info/codesite/1847code.pdf (accessed August 11, 2018).
3. 3. Annette C. Baier, "Why Honesty Is a Hard Virtue," *Reflections on How We Live* (Oxford: Oxford University Press, 2010), p. 109.
4. 4. Henry Sidgwick, *The Methods of Ethics*, 7th ed. (Indianapolis, IN: Hackett, 1907), pp. 315–16. Baier examines honesty constructively from a background of Hume's thought, and Alasdair MacIntyre examines lying in reaction to Kant's and Mill's thought. See Baier, "Why Honesty Is a Hard Virtue"; and MacIntyre, *Ethics and Politics, Selected Essays*, vol. 2 (Cambridge: Cambridge University Press, 2006), chap. 6 (on Mill) and chap. 7 (on Kant).
5. 5. G. J. Warnock, *The Object of Morality* (London: Methuen, 1971), p. 85.
6. 6. See, for example, W. D. Ross, *The Right and the Good* (Oxford: Clarendon, 1930), chap. 2.
7. 7. For a trust-based approach to medicine and nursing, as well as other professional roles, see Terrence M. Kelly, *Professional Ethics: A Trust-Based Approach* (Lanham, MD: Lexington Books, 2018). For a study of the "essential role that trust plays in effective doctor–patient relationships," despite its frequent neglect, see Nicola Brennan, Rebecca Barnes, Mike Calnan, et al., "Trust in the Health-Care Provider–Patient Relationship: A Systematic Mapping Review of The Evidence Base," *International Journal for Quality in Health Care* 25 (2013): 682–88.
8. 8. Cf. Raanan Gillon, "Is There an Important Moral Distinction for Medical Ethics Between Lying and Other Forms of Deception?" *Journal of Medical Ethics* 19 (1993): 131–32; and Jennifer Jackson, *Truth, Trust, and Medicine* (London: Routledge, 2001). A survey of public attitudes toward deception in medicine also found that "while the majority of … respondents opposed outright lying in medical contexts, they were prepared to accept partial disclosure and the use of placebos when it is in the patient's interests or when it is what the person would want." See Jonathan Pugh, Guy Kahane, Hannah Maslen,

and Julian Savulescu, "Lay Attitudes toward Deception in Medicine: Theoretical Considerations and Empirical Evidence," *AJOB Empirical Bioethics* 7, no 1 (2016): 31–38.

9. 9. On different sociocultural contexts of nondisclosure and the need in health care for what is often called "cultural competence," see Antonella Surbone, "Telling the Truth to Patients with Cancer: What Is the Truth?" *Lancet Oncology* 7 (2006): 944–50. See further Loretta M. Kopelman, "Multiculturalism and Truthfulness: Negotiating Difference by Finding Similarities," *South African Journal of Philosophy* 19 (2000): 51–55.
10. 10. Bettina Schöne-Seifert and James F. Childress, "How Much Should the Cancer Patient Know and Decide?" *CA—A Cancer Journal for Physicians* 36 (1986): 85–94.
11. 11. See Donald Oken, "What to Tell Cancer Patients: A Study of Medical Attitudes," *JAMA: Journal of the American Medical Association* 175 (1961): 1120–28; and Dennis H. Novack et al., "Changes in Physicians' Attitudes toward Telling the Cancer Patient," *JAMA: Journal of the American Medical Association* 241 (March 2, 1979): 897–900.
12. 12. N. Horikawa, T. Yamazaki, M. Sagawa, and T. Nagata, "Changes in Disclosure of Information to Cancer Patients in a General Hospital in Japan," *General Hospital Psychiatry* 22 (2000): 37–42. See similar results in T. S. Elwyn, M. D. Fetters, W. Gorenflo, and T. Tsuda, "Cancer Disclosure in Japan: Historical Comparisons, Current Practices," *Social Science and Medicine* 46 (May 1998): 1151–63; and a follow-up study by N. Horikawa, T. Yamazaki, M. Sagawa, and T. Nagata, "The Disclosure of Information to Cancer Patients and Its Relationship to Their Mental State in a Consultation-Liaison Psychiatry Setting in Japan," *General Hospital Psychiatry 21 (September–October 1999): 368–73.*
13. 13. Elisa J. Gordon and Christopher K. Daugherty, "'Hitting You over the Head': Oncologists' Disclosure of Prognosis to Advanced Cancer Patients," *Bioethics* 17 (2003): 142–68; Andrea C. Enzinger, Baohui Zhang, Deborah Schrag, and Holly G. Prigerson, "Outcomes of Prognostic Disclosure: Associations with Prognostic Understanding, Distress, and Relationship with Physician among Patients with Advanced Cancer," *Journal of Clinical Oncology* 33 (2015): 3809–16; Rebecca G. Hagerty, Phyllis N. Butow, Peter M. Ellis, et al., "Communicating with Realism and Hope: Incurable Cancer Patients' Views on the Disclosure of Prognosis," *Journal of Clinical Oncology* 23 (2005): 1278–88.
14. 14. James Boswell, *Life of Johnson*, as quoted in Alan Donagan, *The Theory of Morality* (Chicago: University of Chicago Press, 1997), p. 89.
15. 15. Nicholas A. Christakis, *Death Foretold: Prophecy and Prognosis in Medical Care* (Chicago: University of Chicago Press, 1999), esp. chap. 5. See also G. G. Palmboom, D. L. Willems, N. B. A. T. Janssen, and J. C. J. M. de Haes, "Doctor's Views on Disclosing or Withholding Information on Low Risks of Complication," *Journal of Medical Ethics* 33 (2007): 67–70.
16. 16. Joel Stein, "A Fragile Commodity," *JAMA: Journal of the American Medical Association* 283 (January 19, 2000): 305–6.
17. 17. Thurstan B. Brewin, "Telling the Truth" (Letter), *Lancet* 343 (June 11, 1994): 1512.
18. 18. Antonella Surbone, "Truth Telling to the Patient," *JAMA: Journal of the American Medical Association* 268 (October 7, 1992): 1661–62; and Surbone, Claudia Ritossa, and Antonio G. Spagnolo, "Evolution of Truth-Telling Attitudes and Practices in Italy," *Critical Reviews in Oncology-Hematology* 52 (December 2004): 165–72. See also Baback B. Gabbay et al., "Negotiating End-of-Life Decision Making: A Comparison of Japanese and U. S. Residents' Approaches," *Academic Medicine* 80 (2005): 617–21.
19. 19. Daniel Rayson, "Lisa's Stories," *JAMA: Journal of the American Medical Association* 282 (November 3, 1999): 1605–6.
20. 20. Linda T. Kohn, Janet M. Corrigan, and Molla S. Donaldson, for the Institute of Medicine Committee on Quality of Health Care in America, *To Err Is Human: Building a Safer Health System* (Washington, DC: National Academies Press, 2000). This influential report defines a *medical error* as "the failure of a planned action to be completed as intended (i.e., error of execution) or the use of a wrong plan to achieve an aim (i.e., error of planning)," p. 45.
21. 21. For a critical analysis of "medical harm" and "iatrogenic illness" (literally, illness induced by the physician), see Virginia A. Sharpe and Alan I. Faden, *Medical Harm: Historical, Conceptual, and Ethical Dimensions of Iatrogenic Illness* (Cambridge: Cambridge University Press, 1998).
22. 22. See Robert D. Truog, David M. Browning, Judith A. Johnson, and Thomas H. Gallagher, *Talking with Patients and Families about Medical Error* (Baltimore: Johns Hopkins University Press, 2011), p. vii. The

quotation is from the Foreword by Lucian L. Leape.

23. 23. See Rae M. Lamb et al., "Hospital Disclosure Practices: Results of a National Survey," *Health Affairs* 22 (2003): 73–83; Lamb, "Open Disclosure: The Only Approach to Medical Error," *Quality and Safety in Health Care* 13 (2004): 3–5; Lisa Lehmann et al., "Iatrogenic Events Resulting in Intensive Care Admission: Frequency, Cause, and Disclosure to Patients and Institutions," *American Journal of Medicine* 118 (2005): 409–13; and Allen Kachalia, "Improving Patient Safety through Transparency," *New England Journal of Medicine* 369 (2013): 1677–79.
24. 24. See Thomas H. Gallagher et al., "Patients' and Physicians' Attitudes Regarding the Disclosure of Medical Errors," *JAMA: Journal of the American Medical Association* 289, no. 8 (February 26, 2003): 1001–7; and Gallagher et al., "Choosing Your Words Carefully: How Physicians Would Disclose Harmful Medical Errors to Patients," *Archives of Internal Medicine* 166 (2006): 1585–93. See also David K. Chan et al., "How Surgeons Disclose Medical Errors to Patients: A Study Using Standardized Patients," *Surgery* 138 (November 2005): 851–58.
25. 25. Lisa Iezzoni et al., "Survey Shows That at Least Some Physicians Are Not Always Open or Honest with Patients," *Health Affairs* 31, no. 2 (2012): 383–91.
26. 26. See Steve S. Kraman and Ginny Hamm, "Risk Management: Extreme Honesty May Be the Best Policy," *Annals of Internal Medicine* 131 (December 21, 1999): 963–67; Allen Kachalia et al., "Liability and Costs before and after the Implementation of a Medical Error Disclosure Program," *Annals of Internal Medicine* 153 (2010): 213–21; and Susan D. Moffatt-Bruce, Francis D. Ferdinand, and James I. Fann, "Patient Safety: Disclosure of Medical Errors and Risk Mitigation," *Annals of Thoracic Surgery* 102 (2016): 358–62. See also Nancy Berlinger, *After Harm: Medical Error and the Ethics of Forgiveness* (Baltimore: Johns Hopkins University Press, 2005).
27. 27. Joan Vogel and Richard Delgado, "To Tell the Truth: Physicians' Duty to Disclose Medical Mistakes," *UCLA Law Review* 28 (1980): 55.
28. 28. Truog et al., *Talking with Patients and Families about Medical Error.*
29. 29. The quoted phrase is Joel Feinberg's.
30. 30. Catherine J. Chamberlain et al., "Disclosure of 'Nonharmful' Medical Errors and Other Events: Duty to Disclose," *Archives of Surgery* 147, no. 3 (March 2012): 282–86. Most discussions specifically focus on or assume "harmful medical errors," even if they simply entitle their discussions "medical errors." See, for example, the important study by Thomas H. Gallagher et al., "Patients' and Physicians' Attitudes Regarding the Disclosure of Medical Errors."
31. 31. The Ethics Manual of the American College of Physicians calls on physicians to "disclose to patients information about procedural or judgment errors made in the course of care if such information is material to the patient's well-being." American College of Physicians, *American College of Physicians Ethics Manual*, 6th ed., published in the *Annals of Internal Medicine* 156 (2012): 73–104, and available at https://www.acponline.org/clinical-information/ethics-and-professionalism/acp-ethics-manual-sixth-edition/acp-ethics-manual-sixth-edition (accessed August 15, 2018).
32. 32. In some circumstances, clinicians may have an ethical obligation not only to report colleagues' medical mistakes through appropriate institutional mechanisms but also to disclose these mistakes to patients. See Thomas H. Gallagher, Michelle M. Mello, Wendy Levinson, et al., "Talking with Patients about Other Clinicians' Errors," Sounding Board, *New England Journal of Medicine* 369 (2013): 1752–57.
33. 33. Joanna M. Cain, "Is Deception for Reimbursement in Obstetrics and Gynecology Justified?" *Obstetrics & Gynecology* 82 (September 1993): 475–78.
34. 34. Kaiser Family Foundation, "Survey of Physicians and Nurses," available at http://www.kff.org/1999/1503 (accessed August 20, 2007).
35. 35. Matthew K. Wynia et al., "Physician Manipulation of Reimbursement Rules for Patients: Between a Rock and a Hard Place," *JAMA: Journal of the American Medical Association* 283 (April 12, 2000): 1858–65; and the editorial commentary by M. Gregg Bloche, "Fidelity and Deceit at the Bedside," in the same issue, pp. 1881–84.
36. 36. The case for this assessment is made by Bloche, "Fidelity and Deceit at the Bedside," p. 1883.
37. 37. Victor G. Freeman et al., "Lying for Patients: Physician Deception of Third-Party Payers," *Archives of Internal Medicine* 159 (October 25, 1999): 2263–70.

38. 38. Rachel M. Werner et al., "Lying to Insurance Companies: The Desire to Deceive among Physicians and the Public," *American Journal of Bioethics* 4 (Fall 2004): 53–59, with eleven commentaries on pp. 60–80. See also Dennis H. Novack et al., "Physicians' Attitudes toward Using Deception to Resolve Difficult Ethical Problems," *JAMA: Journal of the American Medical Association* 261 (May 26, 1989): 2980–85. We discuss this study in Chapter 1 (pp. 18–19), where we consider debates about the meaning, range, and scope of deception and lying.
39. 39. Another example comes from restrictive criteria (unchanged for decades) for third-party reimbursement for home nutrition support. See Karen Martin and Carol McGinnis, "Home Nutrition Support: Ethics and Reimbursement," *Nutrition in Clinical Practice* 31, no. 3 (June 2016): 325–33. Beyond possible deception of third-party payers, patients' requests for false or misleading medical diagnoses or certificates arise in many contexts, such as medical excuses from work or school. One ethical analysis focuses on the challenges of "fake diagnoses" or "false or misleading certificates" to benefit patients/clients in four scenarios in Swedish health care: sterilization, asylum, virginity (a cultural requirement for marriage in certain communities), and adoption (from another country). When considerable caution is exercised, some of these measures can be justified in some circumstances. See G. Helgesson and N. Lynöe, "Should Physicians Fake Diagnoses to Help Their Patients?" *Journal of Medical Ethics* 34 (2008): 133–36. In the United States, medical teams have often given a "medical excuse" to reluctant potential living kidney donors, sometimes by providing a false excuse, sometimes by a more general statement about the lack of suitability to be a living donor. Lainie F. Ross defends the latter in "What the Medical Excuse Teaches Us about the Potential Living Donor as Patient," *American Journal of Transplantation* 10, no. 4 (2010): 731–36. In Chapter 9 we discuss such a case in light of different ethical theories.
40. 40. Thomas L. Carson, *Lying and Deception: Theory and Practice* (New York: Oxford University Press, 2010), p. 2.
41. 41. *Griswold v. Connecticut*, 381 U.S. 479 (1965), at 486. On the deep influence of *Griswold* on the development of the right to privacy, see Joanna L. Grossman, "*Griswold v. Connecticut*: The Start of the Revolution," *Verdict*, June 8, 2015, available at https://verdict.justia.com/2015/06/08/griswold-v-connecticut-the-start-of-the-revolution (accessed August 30, 2018).
42. 42. See, for example, Adam D. Moore, *Privacy Rights: Moral and Legal Foundations* (University Park: Pennsylvania State University Press, 2010), p. 5; Michael Katell and Adam D. Moore, "Introduction: The Value of Privacy, Security and Accountability," in *Privacy, Security and Accountability: Ethics, Law and Policy*, ed. Adam D. Moore (London: Rowman & Littlefield International, 2016), p. 3.
43. 43. See Jeffrey M. Skopek, "Reasonable Expectations of Anonymity," *Virginia Law Review* 101 (2015): 691–762.
44. 44. Anita L. Allen, "Genetic Privacy: Emerging Concepts and Values," in *Genetic Secrets: Protecting Privacy and Confidentiality in the Genetic Era*, ed. Mark A. Rothstein (New Haven, CT: Yale University Press, 1997), pp. 31–59. For a wide-ranging account of the meanings and types of "privacies" and their moral and political value, see Allen, *Unpopular Privacy: What Must We Hide* (New York: Oxford University Press, 2011), which includes a defense of "paternalistic privacy" policies. For other examinations of privacy, see Daniel J. Solove, *Understanding Privacy* (Cambridge, MA: Harvard University Press, 2008), which argues that there are multiple forms of privacy, related by virtue of family resemblances; and Helen Nissenbaum, *Privacy in Context: Technology, Policy, and the Integrity of Social Life* (Stanford, CA: Stanford University Press, 2010), which develops an omnibus principle of "contextual integrity" and derives context-relative rights from it on a sector-by-sector basis (see esp. p. 238).
45. 45. In an illuminating historical study, Sarah E. Igo examines "privacy's significance as a cultural sensibility and public value" and shows that it has served as a "catch-all" for a variety of concerns about modern life, social organization, and technology. Igo, *The Known Citizen: A History of Privacy in Modern America* (Cambridge, MA: Harvard University Press, 2018).
46. 46. Charles Fried, "Privacy: A Rational Context," *Yale Law Journal* 77 (1968): 475–93.
47. 47. Warren and Brandeis, "The Right to Privacy," *Harvard Law Review* 4 (1890): 193–220.
48. 48. Thomson, "The Right to Privacy," *Philosophy & Public Affairs* 4, no. 4 (Summer 1975): 295–314, as reprinted in *Philosophical Dimensions of Privacy: An Anthology*, ed. Ferdinand David Schoeman (New York: Cambridge University Press, 1984), pp. 272–89, esp. 280–87. By contrast, Judith Wagner DeCew views privacy as "a multifaceted cluster concept" without deriving it completely from other interests. See

DeCew, *In Pursuit of Privacy: Law, Ethics, and the Rise of Technology* (Ithaca, NY: Cornell University Press, 1997). For a critique of Thomson's theory, see Thomas Scanlon, "Thomson on Privacy," *Philosophy & Public Affairs* 4, no. 4 (Summer 1975): 315–22.

49. 49. James Rachels, "Why Privacy Is Important," p. 292; and Edward Bloustein, "Privacy as an Aspect of Human Dignity," both in *Philosophical Dimensions of Privacy*, ed. Schoeman.
50. 50. See Fried, "Privacy: A Rational Context."
51. 51. Although we consider this argument primary, the consequentialist argument also has considerable merit. These arguments are not mutually exclusive.
52. 52. Joel Feinberg, *Harm to Self*, vol. 3 in *The Moral Limits of the Criminal Law* (New York: Oxford University Press, 1986), chap. 19.
53. 53. On issues in public health ethics, see James F. Childress, Ruth R. Faden, Ruth D. Gaare, et al., "Public Health Ethics: Mapping the Terrain," *Journal of Law, Medicine & Ethics* 30 (2002): 170–78; Madison Powers and Ruth Faden, *Social Justice: The Moral Foundations of Public Health and Health Policy* (New York: Oxford University Press, 2006), especially pp. 80–99; Ronald Bayer, Lawrence O. Gostin, Bruce Jennings, and Bonnie Steinbock, *Public Health Ethics: Theory, Policy, and Practice* (New York: Oxford University Press, 2006); and Ruth Gaare Bernheim, James F. Childress, Richard J. Bonnie, and Alan L. Melnick, *Essentials of Public Health Ethics* (Burlington, MA; Jones and Bartlett Learning, 2015).
54. 54. See James F. Childress, "Surveillance and Public Health Data: The Foundation and Eyes of Public Health," in Bernheim, Childress, Bonnie, and Melnick, *Essentials of Public Health Ethics*, chap. 5; and Lisa M. Lee, Charles M. Heilig, and Angela White, "Ethical Justification for Conducting Public Health Surveillance without Patient Consent," *American Journal of Public Health* 102 (January 2012): 38–44. For thorough and historically grounded analyses of and strong support for public health surveillance, see Amy L. Fairchild, Ronald Bayer, and James Colgrove, *Searching Eyes: Privacy, the State, and Disease Surveillance* (Berkeley: University of California Press, 2007); and Fairchild, Bayer, and Colgrove, "Privacy, Democracy and the Politics of Disease Surveillance," *Public Health Ethics* 1, no. 1 (2008): 30–38.
55. 55. See Wendy K. Mariner, "Mission Creep: Public Health Surveillance and Medical Privacy," *Boston University Law Review* 87 (2007): 347–95.
56. 56. The diabetes program nods to patients' autonomy by allowing them to opt out of all aspects of the program except the registry. See Shadi Chamany et al., "Tracking Diabetes: New York City's A1C Registry," *Milbank Quarterly* 87, no. 3 (2009): 547–70, with responses; and Clarissa G. Barnes, Frederick L. Brancati, and Tiffany L. Gary, "Mandatory Reporting of Noncommunicable Diseases: The Example of the New York City A1C Registry (NYCAR)," *Virtual Mentor, American Medical Association Journal of Ethics* 9 (December 2007): 827–31. For analysis and criticism, see Janlori Goldman et al., "New York City's Initiatives on Diabetes and HIV/AIDS: Implications for Patient Care, Public Health, and Medical Professionalism," *American Journal of Public Health* 98 (May 2008): 16–22.
57. 57. See Lucian V. Torian et al., "Striving toward Comprehensive HIV/AIDS Surveillance: The View from New York City," *Public Health Reports* 122 (2007 Supplement 1): 4–6; and Amy L. Fairchild and Ronald Bayer, "HIV Surveillance, Public Health, and Clinical Medicine: Will the Walls Come Tumbling Down?" *New England Journal of Medicine* 365 (August 25, 2011): 685–87.
58. 58. On the latter point, see Goldman et al., "New York City's Initiatives on Diabetes and HIV/AIDS," p. 17.
59. 59. See Mark A. Rothstein, "Genetic Secrets: A Policy Framework," in *Genetic Secrets*, ed. Rothstein, chap. 23.
60. 60. See Andreas-Holger Maehle, *Contesting Medical Confidentiality: Origins of the Debate in the United States, Britain, and Germany* (Chicago: University of Chicago Press, 2016), which looks at the contested journeys of medical confidentiality in these three countries in the late nineteenth and early twentieth centuries. See also Philip Rieder, Micheline Louis-Courvoisier, and Philippe Huber, "The End of Medical Confidentiality? Patients, Physicians and the State in History," *Medical Humanities* 42 (2016): 149–54, which argues "that medical practices of secrecy were regularly attacked in the past, and that the nature of medical confidentiality evolved through time depending on physicians' values and judgements" and societal pressures.
61. 61. Mark Siegler, "Confidentiality in Medicine—A Decrepit Concept," *New England Journal of Medicine* 307 (1982): 1518–21. See also Bernard Friedland, "Physician–Patient Confidentiality: Time to Re-

examine a Venerable Concept in Light of Contemporary Society and Advances in Medicine," *Journal of Legal Medicine* 15 (1994): 249–77; and, on continued erosion of the concept, Beverly Woodward, "Confidentiality, Consent and Autonomy in the Physician-Patient Relationship," *Health Care Analysis* 9 (2001): 337–51.

62. 62. *Estate of William Behringer v. Medical Center at Princeton*, 249 N.J. Super. 597, 592 A.2d 1251 (1991).
63. 63. Barry D. Weiss, "Confidentiality Expectations of Patients, Physicians, and Medical Students," *JAMA: Journal of the American Medical Association* 247 (1982): 2695–97. According to a later study in the United Kingdom, respondents "believe that few professionals have access to their records but indicated that other people should have access if required." Bolton Research Group, "Patients' Knowledge and Expectations of Confidentiality in Primary Health Care: A Quantitative Study," *British Journal of General Practice* 50 (2000): 901–2.
64. 64. George L. Anesi, "The 'Decrepit Concept' of Confidentiality, 30 Years Later," *Virtual Mentor* 14, no. 9 (2012): 708–11.
65. 65. See the UK General Medical Council, *Confidentiality: Good Practice in Handling Patient Information* (January 2017), 9a and 13–15, available at https://www.gmc-uk.org/-/media/documents/confidentiality-good-practice-in-handling-patient-information—english-0417_pdf-70080105.pdf (accessed August 12, 2018). Based on extensive consultations with the public and practitioners, the process for developing this guidance, which took effect April 25, 2017, offers an excellent model for medical-public engagement on standards for confidentiality. See Ipsos MORI Social Research Institute, *Exploring Patient and Public Attitudes towards Medical Confidentiality: Findings from Discussion Groups and In-depth Interviews*, prepared for the General Medical Council (GMC), April 2016, available at https://www.gmc-uk.org/-/media/documents/Exploring_patient_and_public_attitudes_towards_medical_confidentiality_FINAL_270416.pdf_65939141.pdf (accessed August 12, 2018).
66. 66. *Bratt v. IBM*, 467 N.E.2d 126 (1984); *Bratt, et al. v. IBM*, 785 F.2d 352 (1986).
67. 67. American Psychiatric Association, *Psychiatric Services in Correctional Facilities*, 3rd ed. (Washington, DC: American Psychiatric Association, 2016); Emil R. Pinta, "Decisions to Breach Confidentiality When Prisoners Report Violations of Institutional Rules," *Journal of the American Academy of Psychiatry and the Law* 37 (2009): 150–54.
68. 68. *Tarasoff v. Regents of the University of California*, 17 Cal. 3d 425 (1976); 131 California Reporter 14 (1976). The majority opinion was written by Justice Tobriner; the dissenting opinion was written by Justice Clark.
69. 69. See Kenneth Appelbaum and Paul S. Appelbaum, "The HIV Antibody-Positive Patient," in *Confidentiality versus the Duty to Protect: Foreseeable Harm in the Practice of Psychiatry*, ed. James C. Beck (Washington, DC: American Psychiatry Press, 1990), pp. 127–28.
70. 70. Griffin Sims Edwards, "Doing Their Duty: An Empirical Analysis of the Unintended Effect of *Tarasoff v. Regents* on Homicidal Activity," *Journal of Law and Economics* 57, no. 2 (May 2014): 321–48. He argues, "The original intention of the [Tarasoff-style duty-to-warn] law was to deter dangerous patients from committing heinous crimes, but what may actually have happened was that the law changed the incentives to the patient and the doctor such that the patient has an incentive to withhold homicidal tendencies, and the doctor has an incentive to not explore homicidal tendencies. … The policy implications are simple and fairly easily employed. A change in law to no duty or discretionary duty should decrease homicides" (p. 344). Edwards also provides a chart of duty-to-warn laws in different states in the United States (p. 325).
71. 71. For views to the contrary, see Michael H. Kottow, "Medical Confidentiality: An Intransigent and Absolute Obligation," *Journal of Medical Ethics* 12 (1986): 117–22; and Kenneth Kipnis, "A Defense of Unqualified Medical Confidentiality," *American Journal of Bioethics* 6 (2006): 7–18 (followed by critical commentaries, pp. 19–41).
72. 72. Jean-Pierre Soubrier, "Self-Crash Murder–Suicide: Psychological Autopsy Essay and Questions about the Germanwings Crash," *Crisis: The Journal of Crisis Intervention and Suicide Prevention* 37, no. 6 (2016): 399–401; Romeo Vitelli, "After the Germanwings Crash: Why Did Andreas Lubitz Crash Flight 9525?" *Psychology Today*, February 16, 2017, available at https://www.psychologytoday.com/us/blog/media-spotlight/201702/after-the-germanwings-crash (accessed August 11, 2018).

73. 73. *(Model) Professional Code for Physicians in Germany—MBO-Ä 1997*, article 9, as published in Resolutions of the 121st German Medical Assembly 2018 in Erfurt, available at https://www.bundesaerztekammer.de/fileadmin/user_upload/downloads/pdf-Ordner/MBO/MBO-AE_EN_2018.pdf (accessed August 12, 2018). The official, authorized version of this document is the German original.
74. 74. *Volk v. DeMeerleer, 386 P.3d 254 (Wa. 2016). On the potential conflicts between this decision and the guidance of the AMA Code of Medical Ethics and the annotations by the American Psychiatric Association, see Jennifer L. Piel and Rejoice Opara, "Does Volk v DeMeerleer* Conflict with the AMA Code of Medical Ethics on Breaching Patient Confidentiality to Protect Third Parties?" *AMA Journal of Ethics* 20, no. 1 (2018): 10–18, available at https://journalofethics.ama-assn.org/article/does-volk-v-demeerleer-conflict-ama-code-medical-ethics-breaching-patient-confidentiality (accessed August 13, 2018).
75. 75. Grant Gillett, "AIDS and Confidentiality," *Journal of Applied Philosophy* 4 (1987): 15–20, from which this case study has been adapted.
76. 76. See Suzanne E. Landis, Victor J. Schoenbach, David J. Weber, et al., "Results of a Randomized Trial of Partner Notification in Cases of HIV Infection in North Carolina," *New England Journal of Medicine* 326 (January 9, 1992): 101–6. See also Michael D. Stein et al., "Sexual Ethics: Disclosure of HIV-Positive Status to Partners," *Archives of Internal Medicine* 158 (February 1998): 253–57.
77. 77. Shona Dalai, Cheryl Johnson, Virginia Fonner, et al., "Improving HIV Test Uptake and Case Finding with Assisted Partner Notification Services," *AIDS* 31, no. 13 (August 24, 2017): 1867-76.
78. 78. These provisions are drawn, with some modifications, from World Health Organization, *Guidelines on HIV Self-Testing and Partner Notification: Supplement to Consolidated Guidelines on HIV Testing Services* (Geneva: WHO, December 2016), available at http://apps.who.int/iris/bitstream/handle/10665/251655/9789241549868-eng.pdf?sequence=1&TSPD_101_R0=5b1daef969957fc81e5694e365c2019bv6300000000000000026877ab49ffff000000000000000000000000005b72ff58002d7d6933 (accessed August 14, 2018).
79. 79. *AMA Code of Medical Ethics 2016–2017 Edition*, 8:1, p. 125. This provision was adopted in June 2008, updated in 2010 and again in 2016 (copyright 2018); while its wording has changed somewhat, its content has remained reasonably constant in recent years.
80. 80. World Health Organization, *Guidelines on HIV Self-Testing and Partner Notification.*
81. 81. Sarah Magaziner, Madeline C. Montgomery, Thomas Bertrand, et al., "Public Health Opportunities and Challenges in the Provision of Partner Notification Services: The New England Experience," *BMC Health Services Research* 18, no. 1 (January 31, 2018): 75, available at https://bmchealthservres.biomedcentral.com/track/pdf/10.1186/s12913-018-2890-7 (accessed August 14, 2018).
82. 82. See discussion of this problem, using an absolutist, rule-oriented solution notably different from our on-balance strategy, in Kipnis, "A Defense of Unqualified Medical Confidentiality" (and responses to his claim in the same journal issue).
83. 83. Robert Baker, "Confidentiality in Professional Medical Ethics," *American Journal of Bioethics* 6 (2006): 39–41.
84. 84. Lori B. Andrews et al., eds. (for the Committee on Assessing Genetic Risks, Institute of Medicine), *Assessing Genetic Risks: Implications for Health and Social Policy* (Washington, DC: National Academies Press, 1994), p. 278, and see 264–73. For an argument that medical examiners and forensic pathologists have an ethical (but not a legal) duty to warn a deceased person's relatives about genetics-related risks, such as sudden cardiac death, or about risks from environmental exposure that could affect them too, see Bernice Elger, Katarzyna Michaud, and Patrice Mangin, "When Information Can Save Lives: The Duty to Warn Relatives about Sudden Cardiac Death and Environmental Risks," *Hastings Center Report* 40 (May–June 2010): 39–45.
85. 85. Michael Parker and Anneke Lucassen, "Genetic Information: A Joint Account?" *British Medical Journal* 329 (July 17, 2004): 165–67.
86. 86. This personal-account model is the primary model internationally. See Sandi Dheensa et al., "Health-care Professionals' Responsibility to Patients' Relatives in Genetic Medicine: A Systematic Review and Synthesis of Empirical Research," *Genetics in Medicine* 18 (2016): 290–301.

87. 87. Anneke Lucassen and Roy Gilbar, "Alerting Relatives about Heritable Risks: The Limits of Confidentiality," *BMJ* 361 (April 5, 2018). Available at https://www.bmj.com/content/361/bmj.k1409.full and also at https://www.ncbi.nlm.nih.gov/pmc/articles/PMC5885756/ (accessed August 12 and 31, 2018).
88. 88. Sandi Dheensa, Angela Fenwick, and Anneke Lucassen, "'Is This Knowledge Mine and Nobody Else's? I Don't Feel That.' Patient Views about Consent, Confidentiality and Information-Sharing in Genetic Medicine," *Journal of Medical Ethics* 42 (2016): 174–79. A philosophical view of the relational or connected self that could support this position appears in Heather Widdows, *The Connected Self: The Ethics and Governance of the Genetic Individual*, Cambridge Bioethics and Law (Cambridge: Cambridge University Press, 2013), esp. pp. 81–84. Rather than emphasizing the familial nature of genetic information or special familial obligations, Madison K. Kilbride bases the duty to share potentially helpful genetic information with family members on the general principle of rescue—that is, to help others in need (as we discuss in Chapter 6); by contrast, S. Matthew Liao and Joran MacKenzie stress special obligations in familial relationships. See Kilbride, "Genetic Privacy, Disease Prevention, and the Principle of Rescue"; and Liao and MacKenzie, "Genetic Information, the Principle of Rescue, and Special Obligations," both in the *Hastings Center Report* 48, no. 3 (2018): 10–17, and 18–19, respectively.
89. 89. Agatha M. Gallo, Denise B. Angst, and Kathleen A. Knafl, "Disclosure of Genetic Information within Families: How Nurses Can Facilitate Family Communication," *American Journal of Nursing* 109, no. 4 (2009): 65–69.
90. 90. See Mary Alice Fisher, *The Ethics of Conditional Confidentiality: A Practice Model for Mental Health Professionals* (New York: Oxford University Press, 2013).
91. 91. Paul Ramsey, *The Patient as Person* (New Haven, CT: Yale University Press, 1970), p. xii.
92. 92. See M. Gregg Bloche, "Clinical Loyalties and the Social Purposes of Medicine," *JAMA: Journal of the American Medical Association* 281 (January 20, 1999): 268–74; and Bloche, *The Hippocratic Myth: Why Doctors Are under Pressure to Ration Care, Practice Politics, and Compromise Their Promise to Heal* (New York: Palgrave Macmillan, 2011).
93. 93. See Stephen Toulmin, "Divided Loyalties and Ambiguous Relationships," *Social Science and Medicine* 23 (1986): 784; and Michael D. Robertson and Garry Walter, "Many Faces of the Dual-Role Dilemma in Psychiatric Ethics," *Australian and New Zealand Journal of Psychiatry* 42 (2008): 228–35.
94. 94. See *In re Sampson*, 317 N.Y.S.2d (1970); and *In re McCauley*, 409 Mass. 134, 565 N.E.2d 411 (1991).
95. 95. See broader questions of dual loyalties in military contexts, in comparison with sports medicine, occupational health, and health care in prisons, in Institute of Medicine (now National Academy of Medicine), *Military Medical Ethics: Issues Regarding Dual Loyalties: Workshop Summary* (Washington, DC: National Academies Press, 2009). See also Solomon R. Benatar and Ross E. G. Upshur, "Dual Loyalty of Physicians in the Military and in Civilian Life," *American Journal of Public Health* 98 (December 2008): 2161–67; and Laura Sessums et al., "Ethical Practice under Fire: Deployed Physicians in the Global War on Terrorism," *Military Medicine* 174 (2009): 441–47.
96. 96. In *Oath Betrayed: Torture, Medical Complicity, and the War on Terror* (New York: Random House, 2006), Steven H. Miles asks, "Where were the doctors and nurses at Abu Ghraib?" and challenges medical and health professionals to recognize their responsibilities to disarmed captives and detainees. See also M. Gregg Bloche and Jonathan H. Marks, "When Doctors Go to War," *New England Journal of Medicine* 352, no. 1 (2005): 3–6; Bloche, *The Hippocratic Myth*, chaps. 7 and 8; Michael L. Gross, *Bioethics and Armed Conflict: Moral Dilemmas of Medicine and War* (Cambridge, MA: MIT Press, 2006); and Chiara Lepora and Joseph Millum, "The Tortured Patient: A Medical Dilemma," *Hastings Center Report* 41 (May–June 2011): 38–47.
97. 97. See Curtis Prout and Robert N. Ross, *Care and Punishment: The Dilemmas of Prison Medicine* (Pittsburgh, PA: University of Pittsburgh Press, 1988); Michael Puisis, ed., *Clinical Practice in Correctional Medicine*, 2nd ed. (New York: Mosby, 2006); and Kenneth Kipnis, "Ethical Conflict in Correctional Health Services," in *Conflict of Interest in the Professions*, ed. Michael Davis and Andrew Stark (Oxford: Oxford University Press, 2001), pp. 302–15.
98. 98. In contrast to virtually all of the developed world, the United States maintains capital punishment and has increasingly medicalized it through adoption of lethal injection. The AMA code bans physician participation in executions but does not count some actions as participation, including testifying on the prisoner's medical history, diagnoses, or mental state related to competence to stand trial or on the prisoner's medical diagnoses related to the legal assessment of competence for execution; relieving the

condemned prisoner's acute suffering while awaiting execution; or certifying death after someone else has declared it. See *Code of Medical Ethics of the American Medical Association*, 2016–2017 Edition, 9.7.3, pp. 166–76. The Principles of Medical Ethics and the Opinions of the AMA Council on Ethical and Judicial Affairs make up the *AMA Code of Medical Ethics*. See also Lee Black and Robert M. Sade, "Lethal Injection and Physicians: State Law vs Medical Ethics," *JAMA: Journal of the American Medical Association* 298 (December 19, 2007). For challenges to the AMA's ban, see Lawrence Nelson and Brandon Ashby, "Rethinking the Ethics of Physician Participation in Lethal Injection Execution," *Hastings Center Report* 41 (May–June 2011): 28–37.

99. 99. *Code of Ethics for Nurses with Interpretive Statements* (Silver Spring, MD: American Nurses Association, 2015), p. v.
100. 100. Gregory F. Gramelspacher, Joel D. Howell, and Mark J. Young, "Perceptions of Ethical Problems by Nurses and Doctors," *Archives of Internal Medicine* 146 (March 1986): 577–78. See also R. Walker, S. Miles, C. Stocking, and M. Siegler, "Physicians' and Nurses' Perceptions of Ethics Problems on General Medical Services," *Journal of General Internal Medicine* 6 (1991): 424–29. Similar themes emerged in a Canadian study by Alice Gaudine, Sandra M. LeFort, Marianne Lamb, and Linda Thorne, "Clinical Ethical Conflicts of Nurses and Physicians," *Nursing Ethics* 18 (2011): 9–19.
101. 101. Ann B. Hamric and Leslie J. Blackhall, "Nurse-Physician Perspectives on the Care of Dying Patients in Intensive Care Units: Collaboration, Moral Distress, and Ethical Climate," *Critical Care Medicine* 35 (2007): 422–29. A Canadian study stresses similarities between physicians and nurses in moral commitments and reasoning related to end-of-life care and finds that differences are often a function of hierarchical structures and assigned roles. It concludes that the moral distress experienced by each group could be reduced by "cross disciplinary discussion and mutual recognition of the burden carried by the other." See Kathleen Oberle and Dorothy Hughes, "Doctors' and Nurses' Perceptions of Ethical Problems in End-of-Life Decisions," *Journal of Advanced Nursing* 33, no. 6 (2001): 707–15.
102. 102. See Carol Pavlish, Katherine Brown-Saltzman, Patricia Jakel, and Alyssa Fine, "The Nature of Ethical Conflicts and the Meaning of Moral Community in Oncology Practice," *Oncology Nursing Forum* 41, no. 2 (March 2014): 130–40.
103. 103. Compare our definition with that of a recent report from the Institute of Medicine: A conflict of interest is "a set of circumstances that create a risk that professional judgments or actions regarding a primary interest [such as patient welfare or objective research results] will be unduly influenced by a secondary interest [such as financial gain or a personal relationship]." Institute of Medicine (now National Academy of Medicine), *Conflict of Interest in Medical Research, Education, and Practice*, ed. Bernard Lo and Marilyn J. Field (Washington, DC: National Academies Press, 2009), pp. 45–46, passim. Marc A. Rodwin argues, unpersuasively in our view, that, in contrast to traditional legal concepts of conflict of interest, these definitions create conceptual confusion and will lead to policies that cannot be implemented effectively; Rodwin, "Attempts to Redefine Conflicts of Interest," *Accountability in Research: Policies in Quality Assurance* (December 6, 2017), available at http://www.tandfonline.com/doi/full/10.1080/08989621.2017.1405728 (accessed August 15, 2018).
104. 104. Jean M. Mitchell and T. R. Sass, "Physician Ownership of Ancillary Services: Indirect Demand Inducement or Quality Assurance?" *Journal of Health Economics* 14 (August 1995): 263–89; and, for a study, see Jean M. Mitchell and Jonathan Sunshine, "Consequences of Physicians' Ownership of Health Care Facilities—Joint Ventures in Radiation Therapy," *New England Journal of Medicine* 327 (November 19, 1992): 1497–501; AMA Council on Ethical and Judicial Affairs, "AMA Code of Medical Ethics' Opinions on the Physician as Businessperson," 2013 version, available at https://journalofethics.ama-assn.org/article/ama-code-medical-ethics-opinions-physician-businessperson/2013-02 (accessed August 31, 2018). See, more broadly, the Institute of Medicine, Committee on Conflict of Interest in Medical Research, Education, and Practice, *Conflict of Interest in Medical Research, Education, and Practice*. Chapter 6 focuses on financial interests in physician self-referral and in medical companies whose products physicians prescribe, use, or recommend. This report covers many forms of conflict of interest in medical practice, research, and education discussed in this section.
105. 105. Iezzoni et al., "Survey Shows That at Least Some Physicians Are Not Always Open or Honest with Patients," p. 383.
106. 106. For an examination of efforts in the United States (and Japan and France) to address self-referral, see Marc A. Rodwin, *Conflicts of Interest and the Future of Medicine* (New York: Oxford University Press,

2011), esp. pp. 117–21, 145–47. See also Bruce J. Hillman, "Trying to Regulate Imaging Self-Referral Is Like Playing Whack-A-Mole," *American Journal of Roentgenology* 189 (2007): 267–68.

107. 107. See E. Haavi Morreim, *Balancing Act: The New Medical Ethics of Medicine's New Economics* (Boston: Kluwer Academic, 1991), which has influenced our discussion.
108. 108. Adriane Fugh-Berman and Shahram Ahari, "Following the Script: How Drug Reps Make Friends and Influence Doctors," *PLOS Medicine* 4 (April 2007): 621–25; Jason Dana and George Loewenstein, "A Social Science Perspective on Gifts to Physicians from Industry," *JAMA: Journal of the American Medical Association* 290 (July 9, 2003): 252–55; and Richard F. Adair and Leah R. Holmgren, "Do Drug Samples Influence Resident Prescribing Behavior? A Randomized Trial," *American Journal of Medicine* 118 (2005): 881–84.
109. 109. Dana Katz et al., "All Gifts Large and Small: Toward an Understanding of the Ethics of Pharmaceutical Industry Gift-Giving," *American Journal of Bioethics* 3 (Summer 2003): 39–45, accompanied by commentaries. A summary and analysis of the relevant psychological research appears in Jason Dana, "How Psychological Research Can Inform Policies for Dealing with Conflicts of Interest in Medicine," in Institute of Medicine, *Conflict of Interest in Medical Research, Education, and Practice*, Appendix D, pp. 358–74. The amount of money some parts of industry put into promotion in these ways indicates their conviction that such promotional activities are effective.
110. 110. For several proposals, see Troyen A. Brennan et al., "Health Industry Practices That Create Conflicts of Interest: A Policy Proposal for Academic Medical Centers," *JAMA: Journal of the American Medical Association* 295 (January 25, 2006): 429–33. See also Institute of Medicine, *Conflict of Interest in Medical Research, Education, and Practice*, esp. chaps. 5 and 6.
111. 111. See AAMC Task Force on Financial Conflicts of Interest in Clinical Research, "Protecting Subjects, Preserving Trust, Promoting Progress: Principles and Recommendations for Oversight of an Institution's Financial Interests in Human Subjects Research (I–II)," *Academic Medicine* 78 (2003): 225–45; and Teddy D. Warner and John P. Gluck, "What Do We Really Know about Conflicts of Interest in Biomedical Research?" *Psychopharmacology* 171 (2003): 36–46.
112. 112. See several chapters in part 4, "Clinical Research," in *Conflicts of Interest in Clinical Practice and Research*, ed. Roy G. Spece, Jr., David S. Shimm, and Allen E. Buchanan (New York: Oxford University Press, 1996).
113. 113. See further Institute of Medicine, *Conflict of Interest in Medical Research, Education, and Practice*, chap. 4.
114. 114. National Commission for the Protection of Human Subjects of Biomedical and Behavioral Research, *The Belmont Report: Ethical Principles and Guidelines for the Protection of Human Subjects of Research* (Washington, DC: DHEW Publication OS 78–*0012*), pp. 2–3; Tom L. Beauchamp and Yashar Saghai, "The Historical Foundations of the Research-Practice Distinction in Bioethics," *Theoretical Medicine and Bioethics* 33 (2012): 45–56; and Code of Federal Regulations, Title 45 (Public Welfare), Part 46 (Protection of Human Subjects), sec. 102 (2005), http://www.hhs.gov/ohrp/humansubjects/guidance/45cfr46.html (accessed July 15, 2011).
115. 115. Cf. Franklin G. Miller, "Revisiting the *Belmont Report:* The Ethical Significance of the Distinction between Clinical Research and Medical Care," *APA Newsletter on Philosophy and Medicine* 5 (Spring 2006): 10–14; and Miller and Howard Brody, "The Clinician-Investigator: Unavoidable but Manageable Tension," *Kennedy Institute of Ethics Journal* 13 (2003): 329–46.
116. 116. Nancy E. Kass, Ruth R. Faden, Steven N. Goodman, Peter Pronovost, Sean Tunis, and Tom L. Beauchamp, "The Research-Treatment Distinction: A Problematic Approach for Determining Which Activities Should Have Ethical Oversight," *Hastings Center Report* (Special Report) 43 (2013): S4–S15.
117. 117. Ruth R. Faden, Nancy E. Kass, Steven N. Goodman, Peter Pronovost, Sean Tunis, and Tom L. Beauchamp, "An Ethics Framework for Learning Healthcare Systems: A Departure from Traditional Research Ethics and Clinical Ethics," *Hastings Center Report* (Special Report) 43 (2013): S16–S27.
118. 118. This pledge was formerly known as the "Declaration of Geneva, Physician's Oath," adopted by the General Assembly of the World Medical Association, Geneva, Switzerland, 1948 (when the wording in this particular passage was slightly different) and most recently amended October 2017, available at https://www.wma.net/policies-post/wma-declaration-of-geneva/ (accessed August 14, 2018).
119. 119. As discussed in Chapter 7, pp. 286–90, the research must not exploit populations or participants. Versions of several of these conditions appear in the Nuremberg Code and the US Department of Health

and Human Services, Protection of Human Subjects, 45 CFR 46, HHS regulations, which include four subparts: subpart A, also known as the Federal Policy or the "Common Rule"; subpart B, additional protections for pregnant women, human fetuses, and neonates; subpart C, additional protections for prisoners; and subpart D, additional protections for children, as revised in 2017 with general implementation January 21, 2019. See also Ezekiel J. Emanuel, David Wendler, and Christine Grady, "What Makes Clinical Research Ethical?" *JAMA: Journal of the American Medical Association* 283 (May 24/31, 2000): 2701–11; and James F. Childress, *Priorities in Biomedical Ethics* (Philadelphia: Westminster Press, 1981), chap. 3.

120. 120. Disclosed by surgeon William DeVries at the University of Utah; see Denise Grady, "Summary of Discussion on Ethical Perspectives," in *After Barney Clark: Reflections on the Utah Artificial Heart Program*, ed. Margery W. Shaw (Austin: University of Texas Press, 1984), p. 49.
121. 121. On some basic ethical issues about comparative effectiveness research, see Ruth R. Faden, Tom L. Beauchamp, and Nancy Kass, "Informed Consent, Comparative Effectiveness Research, and Learning Healthcare," *New England Journal of Medicine* 370 (February 20, 2014): 766–68. Although we mainly use the terms *treatment* and *therapy*, our discussion in this section also applies to diagnostic and preventive procedures, among others.
122. 122. See Gunnel Elander and Göran Hermerén, "Placebo Effect and Randomized Clinical Trials," *Theoretical Medicine* 16 (1995): 171–82; and Gerald Logue and Stephen Wear, "A Desperate Solution: Individual Autonomy and the Double-Blind Controlled Experiment," *Journal of Medicine and Philosophy* 20 (1995): 57–64.
123. 123. See Benjamin Freedman, "Equipoise and the Ethics of Clinical Research," *New England Journal of Medicine* 317 (July 16, 1987): 141–45; and Eugene Passamani, "Clinical Trials—Are They Ethical?" *New England Journal of Medicine* 324 (May 30, 1991): 1590–91. Thirty years after Freedman's influential article appeared, the debate about "clinical equipoise" continues. See, for example, the contrasting arguments under the topic "Head to Head: Is the Concept of Clinical Equipoise Still Relevant to Research?" in *BMJ* 359 (December 28, 2017). Spencer Phillips Hey, Alex John London, and Charles Weijer argue that we have no better framework for justifying research involving patient participation. Annette Rid and Franklin Miller argue that it is a mistake to use "clinical equipoise" to try to align the ethics of clinical research with the ethics of clinical practice; they instead propose that clinical research protocols be evaluated according to whether the risks to participants are acceptable in light of the anticipated benefits to them and/or to the society.
124. 124. Franklin G. Miller and Steven Joffe, "Equipoise and the Dilemma of Randomized Clinical Trials," *New England Journal of Medicine* 364 (February 3, 2011): 476–80. These authors argue that "equipoise is fundamentally flawed as a criterion for determining whether a randomized clinical trial is justified," but their argument only supports the claim that equipoise is flawed if interpreted as a *sufficient condition* of the justification of RCTs; the proper claim is that equipoise is only a *necessary condition*. This latter claim is warranted and not fundamentally flawed.
125. 125. Fred Gifford, "So-Called 'Clinical Equipoise' and the Argument from Design," *Journal of Medicine and Philosophy* 32 (2007): 135–50; and Ezekiel Emanuel, W. Bradford Patterson, and Samuel Hellman, "Ethics of Randomized Clinical Trials," *Journal of Clinical Oncology* 16 (1998): 365–71. Some suggest that the fact that patients with advanced cancer have benefited from novel agents that have minimal toxicity raises a challenge for ethical randomized clinical trials. However, rather than abandon the clinical equipoise standard for clinical trials, others argue that "it would be useful to develop new study designs that prove efficacy rapidly, and consensus criteria that determine threshold values for salutary effects beyond which a randomized trial is not necessary," as in Razelle Kurzrock and David J. Stewart, "Equipoise Abandoned? Randomization and Clinical Trials," *Annals of Oncology* 24, no. 10 (October 2013): 2471–74.
126. 126. Don Marquis, "How to Resolve an Ethical Dilemma Concerning Randomized Clinical Trials," *New England Journal of Medicine* 341 (August 26, 1999): 691–93.
127. 127. Jeremy Howick, "Questioning the Methodologic Superiority of 'Placebo' over 'Active' Controlled Trials," *American Journal of Bioethics* 9 (2009): 34–48. See also articles by his critics in this issue and his reply in "Reviewing the Unsubstantiated Claims for the Methodological Superiority of 'Placebo' over 'Active' Controlled Trials: Reply to Open Peer Commentaries," *American Journal of Bioethics* 9 (2009): 5–7. See also Benjamin Freedman, Kathleen Glass, and Charles Weijer, "Placebo Orthodoxy in Clinical

Research II: Ethical, Legal, and Regulatory Myths," *Journal of Law, Medicine & Ethics* 24 (1996): 252–59.

128. 128. Franklin G. Miller, "The Ethics of Placebo-Controlled Trials," in *The Oxford Textbook of Clinical Research Ethics*, ed. Ezekiel Emanuel et al. (New York: Oxford University Press, 2008), pp. 261–72. See also Robert Temple and Susan S. Ellenberg, "Placebo-Controlled Trials and Active-Control Trials in the Evaluation of New Treatments. Part 1: Ethical and Scientific Issues," *Annals of Internal Medicine* 133 (2000): 455–63; and Ellenberg and Temple, "Placebo-Controlled Trials and Active-Control Trials in the Evaluation of New Treatments. Part 2: Practical Issues and Specific Cases," *Annals of Internal Medicine* 133 (2000): 464–70.
129. 129. See M. A. Fischl et al., "The Efficacy of Azidothymidine (AZT) in the Treatment of Patients with AIDS-Related Complex: A Double-Blind, Placebo-Controlled Trial," *New England Journal of Medicine* 317 (1987): 185–91; and D. D. Richman et al., "The Toxicity of Azidothymidine (AZT) in the Treatment of Patients with AIDS and AIDS-Related Complex: A Double-Blind, Placebo-Controlled Trial," *New England Journal of Medicine* 317 (1987): 192–97.
130. 130. This summary is drawn from Thomas B. Freeman et al., "Use of Placebo Surgery in Controlled Trials of a Cellular-Based Therapy for Parkinson's Disease," *New England Journal of Medicine* 341 (September 23, 1999): 988–92.
131. 131. Years later, Scott Y. H. Kim and colleagues conducted semi-structured interviews with a number of participants from three trials for Parkinson's disease that used sham surgery as a control arm. They found that the participants met the standards for informed consent in terms of their understanding of the trial's special features, the rationale for this design, their decision to participate, etc. Hence, they concluded that concerns about informed consent should not be considered "a special ethical barrier to such studies." See Kim et al., "Sham Surgery Controls in Parkinson Disease Clinical Trials: Views of Participants," *Movement Disorders* 27, no. 11 (September 15, 2012): 1461–65.
132. 132. For a variety of ethical views, see Freeman et al., "Use of Placebo Surgery in Controlled Trials of a Cellular-Based Therapy for Parkinson's Disease"; Ruth Macklin, "The Ethical Problems with Sham Surgery in Clinical Research," *New England Journal of Medicine* 341 (September 23, 1999): 992–96; and Franklin G. Miller, "Sham Surgery: An Ethical Analysis," *American Journal of Bioethics* 3 (2003): 41–48, with several commentaries (pp. 50–71). Two decades later, similar clinical trials continued, and some commentators believe that the ethical debate is settled in favor of sham surgery trials for Parkinson's disease as long as certain conditions are met, while others stress the "ethical complexities" that remain. For the former view, see Sophie L. Niemansburg et al., "Reconsidering the Ethics of Sham Interventions in an Era of Emerging Technologies," *Surgery* 157 (2015): 801–10; for the latter view, see Teresa Swift and Richard Huxtable, "The Ethics of Sham Surgery in Parkinson's Disease: Back to the Future?" *Bioethics* 27, no. 4 (2013): 175–85. One survey indicates that the Parkinson disease clinical research community will probably not be convinced of the efficacy of future neurosurgical interventions for this condition unless clinical trials use a sham control. See Scott Y. H. Kim et al., "Science and Ethics of Sham Surgery: A Survey of Parkinson Disease Clinical Researchers," *Archives of Neurology* 62 (September 2005): 1357–60.
133. 133. See Greg Ball, Linda B. Piller, and Michael H. Silverman, "Continuous Safety Monitoring for Randomized Controlled Clinical Trials with Blinded Treatment Information: Part 1: Ethical Considerations," *Contemporary Clinical Trials* 32, Supplement 1 (September 2011): S2–4.
134. 134. For vigorous calls for the independence and integrity of the DSMB, in light of charges of breaches of the "wall" between DSMBs and sponsors, see Jeffrey M. Drazen and Alastair J. J. Wood, "Don't Mess with the DSMB," *New England Journal of Medicine* 363 (July 29, 2010): 477–78; and Catherine D. DeAngelis and Phil B. Fontanarosa, "Ensuring Integrity in Industry-Sponsored Research: Primum Non Nocere, Revisited," *JAMA: Journal of the American Medical Association* 303 (2010): 1196–98.
135. 135. This is Freedman's proposal in "Equipoise and the Ethics of Clinical Research."
136. 136. These conditions and our arguments throughout this section can profitably be compared to the following influential sources: Council for International Organizations of Medical Sciences, *International Ethical Guidelines for Health-Related Research Involving Humans* (Geneva: CIOMS, 2016), available at http://www.cioms.ch (accessed September 3, 2018); National Bioethics Advisory Commission (NBAC), *Ethical and Policy Issues in International Research: Clinical Trials in Developing Countries*, vol. 1 (Bethesda, MD: National Bioethics Advisory Commission, 2001), available at

https://bioethicsarchive.georgetown.edu/nbac/clinical/Vol1.pdf (accessed September 3, 2018); Nuffield Council on Bioethics, *The Ethics of Research Related to Healthcare in Developing Countries* (London: Nuffield Council, 2008), available at http://nuffieldbioethics.org/wp-content/uploads/2014/07/Ethics-of-research-related-to-healthcare-in-developing-countries-I.pdf (accessed September 2, 2018); US Department of Health and Human Services, Food and Drug Administration, Title 21, Code of Federal Regulations, Part 314 (current as of April 1, 2017), available at http://www.accessdata.fda.gov/scripts/cdrh/cfdocs/cfcfr/cfrsearch.cfm (accessed September 2, 2018); International Conference on Harmonisation; *Choice of Control Group and Related Issues in Clinical Trials*, *Federal Register* 66, no. 93, May 14, 2001, 24390–91.

9

Moral Theories

We have mentioned several types of moral theory in earlier chapters without discussing their nature and their value for biomedical ethics. In this chapter we explicate four influential theories: utilitarianism, Kantianism, rights theory, and virtue ethics. Knowledge of these theories is indispensable for reflective study in biomedical ethics because much of the field's literature presumes familiarity with them. Each theory casts light on important aspects of moral thinking in the biological sciences, medicine, nursing, and public health.

Many textbook approaches to moral theory explicate several competing theories and then proceed to criticize them. The criticisms are sometimes so severe that each theory seems fatally wounded, and readers become skeptical about the value of ethical theory in general. Defects and excesses can be found in all major theories, but the theories discussed in this chapter all contain insights and arguments that deserve careful and constructive study. Our goal is to criticize what is questionable or limited in each type of theory and to appropriate what stands to make a contribution to practical ethics. All four types of theory embody significant insights and deserve careful study even though none can justifiably claim to be the only defensible theory.

We sometimes refer to our own account of ethics as a theory, but a word of caution is in order about this term. "Ethical theory" and "moral theory" are commonly used to refer to (1) abstract moral reflection and argument, (2) systematic presentation of the basic components of ethics, (3) an integrated body of moral norms, and (4) a systematic justification of basic moral norms. We attempt to construct a coherent body of virtues, rights, principles, and rules for biomedical ethics. We do not claim to have developed a comprehensive ethical theory in ways suggested by the combination of (3) and (4). We present an organized system of principles and engage in systematic reflection and argument; but we present only some elements of a general moral theory.

Each section of the present chapter (except the first and the last) is divided into subsections structured by (1) an overview of the type of theory under consideration, (2) presentation of how the theory's proponents might approach a difficult case in biomedical ethics, (3) examination of criticisms of the theory's limitations and problems, and (4) a constructive evaluation of the theory's contribution to moral reflection. We recognize the value of various aspects of all four theories discussed in this chapter. We do not, however, hold that the goal of philosophical ethics is to identify the single best theory and give it moral priority. There is no reason to rank one of these four types of theory above the others because much can be learned from each.[1]

Criteria for Assessing Moral Theories

We begin with eight conditions of adequacy for a moral theory. These criteria for theory construction set forth ideal conditions for theories, but not such demanding conditions that no theory could satisfy them. The extent to which all available theories only partially satisfy these conditions will not be our concern. Moreover, some theories may be well suited to some area of the moral life but not to all of it. For example, utilitarianism is often an excellent model for public policy, and rights theory is often the best model for protecting the interests of individuals and groups against unjust demands that serve dominant community interests.

Eight conditions provide the basic criteria for assessing ethical theories:[2]

1. 1. *Clarity.* Taken as a whole or in its parts, a theory should be as clear as possible. Although, as Aristotle suggested, we can expect only as much clarity and precision of language as the subject matter allows, more obscurity and vagueness currently exist in various parts of ethical theory relevant to biomedical ethics than ideally should be present.
2. 2. *Coherence.* An ethical theory should be internally coherent. There should be neither conceptual inconsistencies (e.g., "hard medical paternalism is justified only by consent of the patient") nor apparently contradictory statements (e.g., "being virtuous is a moral obligation, but virtuous acts are not obligatory").

If an account contains implications that are incoherent with other parts of the account, some aspect of the theory must be changed in a way that does not produce further incoherence. In Chapter 10 we argue that a major goal of a theory should be to bring into coherence all of its normative elements (principles, virtues, rights, considered judgments, and the like).

3. 3. *Comprehensiveness.* A theory should be as comprehensive as possible. It would be fully comprehensive if it could account for all justifiable moral norms and judgments. Although the principles presented in this book under the headings of respect for autonomy, nonmaleficence, beneficence, and justice are far from a complete system for general normative ethics, they provide a comprehensive general framework for the domain of biomedical ethics. We do not need additional general principles for this purpose, but we do specify these four principles to generate rules such as promise-keeping, truthfulness, privacy, and confidentiality (see Chapter 8). Specified rules stand to increase a theory's comprehensiveness.
4. 4. *Simplicity.* A theory that distills the demands of morality to a few basic norms is preferable to a theory with more norms but no additional and needed content. A theory should have no more norms than are necessary (simplicity in the sense of theoretical parsimony), and also no more norms than people can use without confusion (practical simplicity). However, morality is complicated both theoretically and practically, and a comprehensive moral theory is certain to be complex. If the inherent complexity of morality demands a theory of morality too difficult for practical use, the theory cannot be faulted for this reason alone.
5. 5. *Explanatory power.* A theory has explanatory power when it provides enough insight to help us understand morality: its purpose, its objective or subjective status, how rights are related to obligations, and the like. For the sake of clarity, we should distinguish between normative theories and metaethical theories. A general normative theory should not be held to the task of shedding light on metaethical questions, but the ideal theory is one that seamlessly constructs a normative system while addressing relevant metaethical questions.
6. 6. *Justificatory power.* A theory should provide grounds for justified belief, not merely a reformulation of beliefs we already possess. For example, the distinction between acts and omissions underlies several traditional beliefs in biomedical ethics, such as the belief that killing is impermissible and allowing to die permissible. But a moral theory would be impoverished if it only incorporated this distinction without determining whether and, if so, when the distinction is justifiable. A good theory also should have the power to criticize defective beliefs, no matter how widely accepted the beliefs are.
7. 7. *Output power.* A theory has output power when it produces judgments that were not in the original database of considered moral judgments on which the theory was constructed. If a normative theory merely repeats the judgments regarded as sound prior to the construction of the theory, it would have accomplished nothing. For example, if the parts of a theory pertaining to obligations of beneficence do not yield new judgments about role obligations of care in medicine beyond those assumed in constructing the theory, the theory will amount to no more than a classification scheme. A theory, then, must generate more than a list of axioms present in pretheoretic belief.
8. 8. *Practicability.* A moral theory is unacceptable if its practical requirements are so demanding that they cannot be satisfied or could be satisfied by only a few extraordinary persons or communities. A theory that presents utopian ideals or unfeasible recommendations fails the criterion of practicability. For example, if a theory proposed such high requirements for personal autonomy (see Chapter 4) or such lofty standards of social justice (see Chapter 7) that no person could be autonomous or no society just, the proposed theory would be problematic.

Other criteria of theory construction could be mentioned, but the eight we have identified are the most important for our purposes. A theory can receive a high score on the basis of one or more of these criteria and a low score on the basis of other criteria. For example, utilitarianism is arguably an internally coherent, simple, and comprehensive theory with exceptional output power, but it may not be coherent with some vital considered judgments, especially with certain judgments about justice, human rights, and the importance of personal projects. By contrast, Kantian theories are consistent with many of our considered judgments, but their simplicity and output power may be limited.

UTILITARIAN THEORY

Consequentialism is a label affixed to theories holding that actions are right or wrong according to the balance of their good and bad consequences. This term denotes theories that use the promotion of value to determine the rightness or wrongness of actions.

The most prominent consequentialist theory, utilitarianism, concentrates on the value of well-being, which has been analyzed in terms of pleasure, happiness, welfare, preference satisfaction, and the like. Utilitarianism accepts one, and only one, basic principle of ethics: the principle of utility. This principle asserts that we ought always to produce the maximal balance of positive value over disvalue—or the least possible disvalue, if only undesirable results can be achieved. This principle is often formulated as a requirement to do the greatest good for the greatest number, as determined from an impartial perspective that gives equal weight to the legitimate interests of each affected party. The classical origins of this theory are found in the writings of Jeremy Bentham (1748–1832) and John Stuart Mill (1806–1873).

The model bequeathed to philosophy by these authors, principally by Mill, renders utilitarian theory consequentialist, welfarist, aggregative, maximizing, and impartial. It is consequentialist because the moral rightness and obligatoriness of actions are established by their results, and it is welfarist because the rightness of actions is determined by good welfare outcomes. It is impersonal and aggregative because a judgment about right or obligatory action depends on an impartial appraisal that sums up the effects of different possible actions on the welfare of all affected parties.

The Concept of Utility

Utilitarians share the conviction that we should morally assess human actions in terms of their production of maximal value, but they often disagree about which values should be maximized. Many utilitarians maintain that we ought to produce *agent-neutral* or *intrinsic* goods, that is, goods such as happiness, freedom, and health that every rational person values.[3] These goods are valuable in themselves, without reference to their further consequences or to the particular preferences of individuals.

Bentham and Mill are *hedonistic* utilitarians because they conceive utility in terms of happiness or pleasure, two broad terms they treat as synonymous.[4] They acknowledge that many human actions do not appear to be performed for the sake of happiness. For example, when highly motivated professionals, such as research scientists, work themselves to the point of exhaustion in search of new knowledge, they do not appear to be seeking personal happiness. Mill argues that persons of this description are initially motivated by success, recognition, or money, which all promise happiness. Along the way, either the pursuit of knowledge provides happiness or such persons never stop associating their hard work with the success, recognition, or money that they hope to gain.

Various recent utilitarians, by contrast to Mill, argue that a diverse set of values beyond happiness contribute to well-being. Examples are beauty, knowledge, health, success, understanding, enjoyment, and deep personal relationships.[5] Even when their lists differ, these utilitarians concur that we should assess the greatest good of an act in terms of the total intrinsic value it produces. Other utilitarians hold that the concept of utility does not refer to intrinsic goods but to individuals' preferences—that is, the maximization of the overall satisfaction of the preferences of the individuals affected.

A Case of Risk and Truthfulness

Our analysis of each type of ethical theory considers how its proponents might approach the same case. This case centers on a five-year-old girl who has progressive renal failure and is not responding well on chronic renal dialysis. The medical staff is considering a renal transplant, but its effectiveness is judged questionable. Nevertheless, a "clear possibility" exists that a transplanted kidney will not be affected by the disease process. The parents concur with the plan to try a transplant, but an additional obstacle emerges. The tissue typing indicates that it would be difficult to find a match for the girl. The staff excludes her two siblings, ages two and

four, as too young to provide a kidney. The mother is not histocompatible, but the father is compatible and has "anatomically favorable circulation for transplantation."

While meeting in private with the father, the nephrologist presents the results and indicates that the prognosis for his daughter is "quite uncertain." After reflection, the father decides that he will not donate a kidney to his daughter. His several reasons include his fear of the surgery and lack of "courage," the uncertain prognosis for his daughter even with a transplant, the slight prospect of a cadaver kidney, and the suffering his daughter has already sustained. The father then requests that the physician tell all members of the family that he is not histocompatible. He is afraid that if family members know the truth, they will accuse him of failing to save his daughter when he could do so. He argues that telling the truth would have the effect of wrecking the family. The physician feels "very uncomfortable" with this request but agrees to tell the man's wife that the father should not donate a kidney "for medical reasons."[6]

Utilitarians would evaluate this case primarily in terms of the probable consequences of the different courses of action open to the father and the physician. The goal is to realize the greatest good by balancing the interests of all affected persons. This evaluation depends on judgments concerning probable outcomes. Whether the father ought to donate his kidney depends on the probability of successful transplantation as well as the risks and other costs to him and indirectly to other dependent members of the family. The potential effectiveness of transplantation is questionable and the prognosis for the daughter uncertain, although a possibility exists that a transplanted kidney would not undergo the same disease process. There is only a slight possibility that a cadaver kidney could be obtained.

The girl will probably die without a transplant from either a deceased or a living donor, but the transplant also offers only a small chance of survival. The risk of death to the father from anesthesia during kidney removal is 1 in 10,000 to 15,000 (at the time this case was published). The chance of success with a transplant is likely greater than the probability that the father will suffer a significant harm. Under these circumstances a utilitarian might hold that the father or anyone else similarly situated is obligated to undertake what others might consider a heroic act that surpasses obligation. Given the balance of probable benefits and risks, an uncompromising utilitarian might suggest tissue typing the patient's two siblings and then removing a kidney from one if there were a good match and parental approval. However, utilitarians frequently disagree among themselves about which judgment to make in specific cases because of their different theories of value and their different predictions and assessments of probable outcomes.

Probabilistic judgments would likewise play a role in a physician's utilitarian calculation of the right action in response to the father's request. The physician would take into account a variety of considerations, including whether full disclosure would wreck the family, whether lying to the family would have serious negative effects, and whether the father would subsequently experience serious guilt as a result of his refusal to donate. Utilitarians would argue that the physician is obligated to consider the whole range of facts and possible consequences in light of the best available information.

Act and Rule Utilitarianism

The principle of utility is the ultimate standard of right and wrong for all utilitarians. The influence of this principle in bioethics has been deep and abiding.[7] Controversy has arisen, however, over whether this principle justifies particular acts in particular circumstances or instead justifies general rules that determine which acts are right and wrong. The *rule utilitarian* holds that particular acts and judgments are morally justified by impartially formulated rules that maximize value in a society that adopts them, whereas the *act utilitarian* disregards the level of rules and justifies actions by appealing directly to the principle of utility, as the following chart indicates:

Rule Utilitarianism	*Act Utilitarianism*
Principle of Utility	Principle of Utility
↑	

Moral Rules ↑

↑

Particular Judgments Particular Judgments

The act utilitarian asks, "Which good and bad consequences will probably result from this action in this circumstance?" Although moral rules are useful in guiding human actions, they are also expendable if they do not maximally promote utility in a particular context. For a rule utilitarian, by contrast, an act's conformity to a rule that is justified by utility makes the act right, and the rule is not expendable in a particular context even if following the rule does not maximize utility in that context. Each theory of rule utilitarianism can justify not only basic moral rules, but also moral rights, professional duties, and the like.[8]

Physician Worthington Hooker, a prominent nineteenth-century figure in academic medicine and medical ethics, was an incipient rule utilitarian who attended to rules of truth-telling in medicine as follows:

> The good, which may be done by deception in a *few* cases, is almost as nothing, compared with the evil which it does in *many*, when the prospect of its doing good was just as promising as it was in those in which it succeeded. And when we add to this the evil which would result from a *general* adoption of a system of deception, the importance of a strict adherence to the truth in our intercourse with the sick, even on the ground of expediency, becomes incalculably great.[9]

Hooker argued that widespread deception and other compromises with truth-telling in medicine will have an increasingly negative effect over time and will eventually produce more harm than good.

Act utilitarians, by contrast, argue that observing a rule such as truth-telling does not always maximize the general good, and that such rules are only rough guidelines. They regard rule utilitarians as unfaithful to the fundamental demand of the principle of utility, which is to "maximize value."[10] From this perspective, physicians are not obligated to always tell the truth to their patients or their families. For example, sometimes physicians should lie to instill hope and should appreciate that selective adherence to rules rarely erodes moral rules or general respect for medical morality.

Because of the benefits to society of the general observance of morally justified rules, the rule utilitarian does not abandon rules, even in difficult situations. Abandonment threatens the integrity and existence of both the particular rules and the whole system of rules.[11] The act utilitarian's reply is that although rules such as promise-keeping usually should be faithfully observed to maintain trust, they may be set aside when doing so would maximize overall good.[12]

An Absolute Principle and Its Nonabsolute Rules

From the utilitarian's perspective, the principle of utility is the sole and absolute principle of ethics, but no rule derived from this principle is absolute or unrevisable. For example, rules in medicine against actively ending a patient's life may be overturned or substantially revised, depending on the consequences of having or not having the rules. In Chapter 5 we assessed current debates about whether seriously suffering patients should, at their request, be actively assisted in dying rather than merely allowed to die. The rule utilitarian view is that we should support rules that permit physicians to hasten death if and only if those rules would produce the most utility. Likewise, there should be rules against physician-assisted death if and only if those rules would maximize utility. Utilitarians often point out that many people do not currently support allowing physicians to actively bring about a patient's death because of the adverse social consequences they believe would follow for those directly and indirectly affected. If, however, under a different set of social conditions, legalization of physician-assisted death would maximize overall social welfare, the utilitarian would not prohibit the practice and would find good reason to support it. Utilitarians regard their theory as ever responsive in constructive ways to the need for changes in social practices.

A Critical Evaluation of Utilitarianism

Utilitarianism is an attractive moral theory for the formation of public and institutional policies. However, it is not a fully adequate moral theory even for those areas, much less for all areas of the moral life, for the reasons discussed in this section.

Problems about immoral preferences and actions. Problems arise for utilitarians who are concerned about the maximization of individual preferences when some of these individuals have what our considered moral judgments (see Chapter 10, pp. 439–45) tell us are morally unacceptable preferences. For example, if a research investigator derives great satisfaction from inflicting pain on animals in research or on human subjects in experiments, we would condemn this preference and would seek to prevent it from being satisfied. A theory based on subjective preferences is a plausible theory only if we can formulate a range of *acceptable* preferences and determine acceptability independently of the particular preferences agents happen to have. This task is inconsistent with a pure preference approach to utility, because there is no pure utilitarian way to elevate one set of preferences over another.[13]

A related problem arises from the utility of immoral actions. Suppose the only way to achieve the maximal utilitarian outcome is to perform an immoral act such as killing terminally ill persons to distribute their organs to several others who will die without them. Act utilitarian theory suggests not only that such killing is permissible, but that it is morally obligatory if the killing would in fact achieve an overall maximization of utility.

Overdemandingness: Does utilitarianism require too much? Some forms of utilitarianism demand too much in the moral life, because the principle of utility requires maximizing value. Utilitarians therefore have a difficult time maintaining the distinction between morally obligatory actions and supererogatory actions (see Chapter 2, pp. 46–48). Alan Donagan presents a variety of situations in which utilitarian theory regards an action as obligatory even though the action is ideal and praiseworthy rather than obligatory.[14] For example, Donagan would regard the "voluntary" suicide of frail elderly persons who suffer from severe disabilities and are no longer useful to society as an example of acts that cannot rightly be considered obligatory, regardless of the consequences. The same holds for gifts of bodily parts, such as kidneys and hearts, to save another person's life. If utilitarianism makes such actions obligatory, it is a defective theory.

Bernard Williams and John Mackie offer extensions of this thesis that utilitarianism demands too much. Williams argues that utilitarianism abrades personal integrity by making persons morally responsible for consequences that they *fail to prevent* as much as for those outcomes they *directly cause*, even when the consequences are not of their doing. Mackie argues that the utilitarian's test of right actions is so distant from our moral experience that it is "the ethics of fantasy," because it demands that people strip themselves of goals and relationships they greatly value in order to maximize outcomes for others.[15]

Problems of unjust distributions. Some utilitarian theories in principle permit the interests of the majority in society to override the rights of minorities and do not have the resources to adequately guard against unjust social distributions. The problem is that utilitarians assign no independent weight to rights and justice and their arguments seem indifferent to unjust distributions because they distribute value according to net aggregate satisfaction.[16] Accordingly, if an already prosperous group of persons could have more value added to their lives than could be added to the lives of the indigent in society, the utilitarian must recommend that the added value go to the prosperous group.

An example of problematic, although on balance perhaps justified, distribution appears in the following case. Two researchers set out to determine the most cost-effective way to control hypertension in the American population. They discovered that it is more cost-effective to target patients already being treated for hypertension than to identify new cases of hypertension among persons without regular access to medical care. They concluded that "a community with limited resources would probably do better to concentrate its efforts on improving adherence of known hypertensives (that is, those already identified as sufferers of hypertension), even at a sacrifice in terms of the numbers screened." No other policy would work as efficiently to maximize social

utility as targeting known hypertensives already in contact with physicians. However, this recommendation would exclude the poorest sector of the population with the most pressing need for medical attention from the benefits of publicly funded high blood pressure education and management.[17] (See further our discussion of cost-effectiveness analysis in Chapter 6, pp. 251–52.)

A Constructive Evaluation of Utilitarianism

Despite these criticisms, utilitarianism has many strengths, two of which we have mentioned in other chapters. The first is the significant role the principle of utility can play in formulating public and institutional policies. The utilitarian's requirements for an objective assessment of everyone's interests and an impartial choice to maximize good outcomes for all affected parties are acceptable, indeed, worthy, norms of public policy, except when they might lead to unjust distributions and the like. Second, in our formulation of principles of beneficence in Chapter 6 (esp. pp. 217–24), utility plays an important role. In the present chapter we have characterized utilitarianism as primarily a consequence-based theory, but it is also beneficence based. That is, the theory sees morality primarily in terms of the legitimate goal of promoting welfare and takes that role with appropriate seriousness. As argued previously, nonmaleficence and beneficence are among the most basic of moral principles in biomedical ethics, and utilitarianism is erected at its foundations on these principles.[18]

KANTIAN THEORY

A second type of theory denies much that utilitarian theories affirm. Often called *deontological*[19] and *nonconsequentialist*[20] (that is, a theory of duty holding that some features of actions other than or in addition to consequences make actions right or wrong), this type of theory is now frequently called *Kantian*, because the philosophy of Immanuel Kant (1724–1804) has most penetratingly shaped many of its contemporary formulations.

Consider how a Kantian might approach the previously mentioned case of the five-year-old in need of a kidney. A Kantian presumably would maintain that we should rest our moral judgments on reasons that apply to all persons who are similarly situated. If the father has no generalizable moral obligation to his daughter, then no basis is available for morally criticizing him for not donating a kidney. The strict Kantian maintains that if the father chooses to donate out of affection, compassion, or concern for his dying daughter, his act would lack moral worth, because it would not be based on a generalizable obligation; but the donation would have moral worth if done from the *duty* (or *obligation*) of beneficence. Using one of the girl's younger siblings as a source of a kidney would be morally prohibited because this recourse to children who are too young to consent to donation involves using persons merely as means to others' ends. This principle also precludes coercing the father to donate against his will.

Regarding the father's request for the physician to deceive the family, a strict Kantian views lying as an act that cannot consistently be universalized as a moral norm. The physician should not lie to the man's wife or to other members of the family even if it would help keep the family intact (a consequentialist appeal). Although the physician's statement that the father should not donate a kidney "for medical reasons" is not, on one interpretation, a lie, the physician still is intentionally concealing relevant facts from the wife, an act a Kantian would typically view as morally unacceptable.

A Kantian will also consider whether the rule of confidentiality has independent moral weight, whether the tests the father underwent with the nephrologist established a relationship of confidentiality, and whether the rule of confidentiality protects information about the father's histocompatibility and his reasons for not donating. If confidentiality prohibits the nephrologist from letting the family know that the father is histocompatible, then the Kantian must face an apparent conflict of obligations: truthful disclosure in conflict with confidentiality.

We need to have the rudiments of Kantian theory before us in order to address this conflict.

Obligation Generated by Categorical Rules

In Kant's theory, morality is grounded in reason, rather than mere tradition, intuition, or attitudes such as sympathy. Human beings are creatures with rational powers that motivate them morally, that help them resist tempting desires, and that allow them to prescribe moral rules to themselves. Kant claims that the moral worth of an individual's action depends exclusively on the moral acceptability of the "maxim" (i.e., the general rule of conduct) on which the person is acting. True moral obligation is based on a universally valid rule that determines the individual's will and justifies the action.[21]

For Kant, one must act not only in accordance with, but also for the sake of, the obligation stated in the rule. That is, to have moral worth, a person's motive for acting must come from a recognition that he or she intends that which is morally required. For example, if an employer discloses a health hazard to an employee only because the employer fears a lawsuit, and not because of the importance of truth-telling, the employer has performed the right action but deserves no moral credit for performing it. If agents do what is morally right simply because they are scared, because they derive pleasure from doing that kind of act, or because they seek recognition, they lack the right kind of motive, which is acting for the sake of obligation.

Kant presents the case of a man who desperately needs money and knows that he will not be able to borrow it unless he promises repayment in a definite time, but also knows that he will not be able to repay it within this period. He decides to make a promise that he knows he will break. Kant asks us to examine the man's reason, that is, maxim: "When I think myself in want of money, I will borrow money and promise to pay it back, although I know that I cannot do so." Kant maintains that this maxim cannot pass a test he calls the *categorical imperative.* This imperative tells us what must be done irrespective of our personal desires or goals. In its major formulation, Kant states the categorical imperative as, "I ought never to act except in such a way that I can also will that my maxim become a universal law." Kant says that this general principle justifies all particular imperatives of obligation (all "ought" statements that morally obligate).[22]

The categorical imperative is a canon of the acceptability of moral rules, that is, a criterion for judging the acceptability of the maxims on which we act. This imperative adds nothing to a maxim's content. It simply determines which maxims are objective and valid. The categorical imperative functions by testing what Kant calls the "consistency of maxims": A maxim must be capable of being conceived and willed without contradiction. When we examine the maxim of the person who deceitfully promises, we discover, Kant maintains, that this maxim is incapable of being conceived and willed universally without yielding a contradiction. It is inconsistent with what it presupposes, as if to say, "My promise can be deceitful, though promising cannot be deceitful." The universalized maxim that a deceitful promise is permissible is inconsistent with the institution of promising it presupposes, and this universalized maxim would be undermined if everyone acted on it. Lying, too, works only if the person to whom the lie is told expects or presupposes that people are truthful, but, if universalized, a maxim approving lying would make the purpose of truth-telling impossible, and no one would believe the person who told a lie.[23]

Kant has more than one version or formulation of the categorical imperative. His second formulation is widely cited in biomedical ethics and is far more influential in this field than the first: "One must act to treat every person as an end and never as a means only."[24] It is often stated that this principle categorically requires that we should never treat another as a means to our ends, but this interpretation misrepresents Kant's views. He argues only that we must not treat another *merely* or *exclusively* as a means to our ends. When human research subjects volunteer to test new drugs, they are treated as a means to others' ends, but they have a choice in the matter and retain control over their lives. Kant does not prohibit such uses of consenting persons. He insists only that they be treated with the respect and moral dignity to which every person is entitled.

Autonomy and Heteronomy

We saw in Chapter 4 that the word *autonomy* typically refers to what makes judgments and actions one's own. Kant's theory of autonomy is significantly different: Persons have "autonomy of the will" if and only if they

knowingly act in accordance with the universally valid moral principles that pass the requirements of the categorical imperative. Kant contrasts this moral autonomy with "heteronomy," which refers to any determinative influence over the will other than motivation by valid moral principles.[25] If, for example, people act from passion, desire, personal ambition, or self-interest, they act heteronomously. Only a rational will acting morally chooses autonomously. Kant regards acting from fear, pity, impulse, personal projects, and habit as heteronomous, as are actions manipulated or coerced by others.

To say that an individual must *accept* a moral principle to qualify as autonomous does not mean that the principle is subjective or that individuals must create (author or originate) their moral principles. Kant requires only that each individual will the acceptance of valid moral principles. If a person freely accepts objective moral principles, that person is a moral lawgiver unto himself or herself. In a bold account, Kant extends beyond the nature of autonomy to its value: "The principle of autonomy" is "the sole principle of morals" and autonomy gives people respect, value, and proper motivation. A person's dignity—indeed, "sublimity"—comes from being morally autonomous.[26]

Kant's theory of autonomy is not about respect for the self-determination of agents who make judgments and set personal goals. It is exclusively about *moral* self-determination. Nonetheless, Kant's second formulation of the categorical imperative is close in important respects to the normative commitments in the principle of respect for autonomy we developed in Chapter 4.

Contemporary Kantian Ethics

Several writers in ethical theory have accepted and developed influential Kantian moral theories, broadly construed.

An example is *The Theory of Morality* by Alan Donagan. He seeks the "philosophical core" of the morality expressed in the Hebrew-Christian tradition, which he interprets in secular rather than religious terms. Donagan's account relies heavily on Kant's theory of persons as ends in themselves, especially the imperative that one must treat humanity as an end and never as a means only. Donagan expresses the fundamental principle of the Hebrew-Christian tradition as a Kantian principle grounded in rationality: "It is impermissible not to respect every human being, oneself or any other, as a rational creature."[27]

A second Kantian theory derives from the work of John Rawls, who challenged utilitarian theories by developing Kantian themes of reason, autonomy, individual worth, self-respect, and equality.[28] His book *A Theory of Justice* uses Kant's moral theory to construct the foundation of a theory of justice (which we treat in Chapter 7, pp. 274–76). For Rawls, the right to individual autonomy of an agent (as discussed in Chapter 4, pp. 99–106) does not outweigh what rational moral principles demand. Even conscientious acts of individual autonomy do not merit respect unless they are in accord with moral principles.[29]

Several philosophers, including Bernard Williams and Thomas Nagel, present views about "deontological constraints" that are related to Kant's injunction never to use another person merely as a means.[30] They interpret Kant as correctly maintaining that certain actions are impermissible regardless of the consequences. For example, in research involving human subjects, even if achieving great breakthroughs would have good consequences for millions of people, researchers would be treating their subjects unethically if they violated fundamental ethical constraints such as failing to obtain subjects' (or their surrogates') voluntary, informed consent.

Another influential Kantian, Christine Korsgaard, warns that we stand to misconstrue Kant if we interpret his moral theory through the lens of such constraints. She argues that when philosophers contrast utilitarian and Kantian theories, they often miss the fact that these two types of moral theory take strikingly different views about the subject matter of ethics. Whereas utilitarians take the subject matter to be the outcomes of actions, Kantians see the subject matter as the quality of relationships, what we owe to others, and the like. Utilitarians hold that one should be just and beneficent in relationships to others because it maximizes the good, but in

Kantian theory the norm that one should produce good outcomes derives from norms of proper relationships. Korsgaard argues that it is a mistake to present Kantian theory as defending deontological constraints as if they were constraints on the goal of promoting the good. In her interpretation, Kant does not support the claim that there is a general duty to promote the good that then must be constrained.[31]

Another Kantian philosopher, Onora O'Neill, has extended Kantian thought into several areas of biomedical ethics, public health, and global justice. Her themes focus heavily on "principled autonomy," public reason, a robust interpretation of universalizability, and the importance of creating conditions of trust.[32] O'Neill's Kantian views are discussed in Chapter 4 (p. 119).

A Critical Evaluation of Kantian Theory

Like utilitarianism, Kant's theory and modern reformulations do not provide a fully convincing or adequately comprehensive theory of the moral life.

The problem of conflicting obligations. Kant construes moral requirements as categorical imperatives, but this theory is inadequate to handle the problem of conflicting obligations. Suppose we have promised to take our children on a long-anticipated trip, but now find that if we do so, we cannot assist our sick mother in the hospital. A rule of promise-keeping in conflict with an obligation of care generates this conflict. Conflict can also arise from a single moral rule rather than from two different rules, as, for example, when two promises come into conflict, although the promisor could not have anticipated the conflict when making the promises. Because moral rules are *categorical* imperatives in Kant's theory, he is committed to the view that we are obligated to perform both actions. Any ethical theory that winds up with this conclusion is unsatisfactory.[33]

Overemphasizing law while underemphasizing relationships. Kant's arguments concentrate on obligations from moral law, and some recent Kantian theories feature a contractual basis for obligations. But whether contract, moral law, and related staples of Kantianism deserve to occupy this central position in a moral theory is questionable. These visions of the moral life fail to capture much that is morally important in personal relationships. For example, we rarely think or act in terms of law, contract, or absolute rules in relationships among friends and family.[34] This feature of the moral life suggests that Kant's theory (like utilitarianism) is better suited for relationships among strangers than for relationships among friends or other intimates, including patients and research subjects.

Virtue, emotion, and moral worth. Kant maintains that actions done from sympathy, emotion, and the like have no moral worth; only actions performed from duty (i.e., the motive of duty) have moral worth. Kant does not disallow or even discourage sympathy and moral emotions, but these motives count for nothing morally. Yet, as we argued in Chapter 2 (pp. 31–33), actions done from sympathy, emotion, and the like do often have moral worth, at least under some conditions. Persons with appropriate feelings and concern about their friends, for example, are morally worthier than persons who discharge obligations of friendship entirely from a sense of duty. Of course, we want people to be attentive to and to discharge their obligations, and there is nothing wrong with a motive of duty; but motivation from deep care and concern can also be meritorious.[35]

A Constructive Evaluation of Kantian Theory

Kant argues that when good reasons support a moral judgment, those reasons are good for all relevantly similar circumstances. Most moral theories now accept roughly this claim, and Kant must be credited for a compelling and far-reaching theoretical account. For example, if we are required to obtain valid consent from all human subjects of biomedical research, we cannot make exceptions of certain persons merely because we could advance science by doing so. We cannot use institutionalized populations without consent any more than we can use people who are not in institutions without their consent. Kant and many Kantians have discerningly driven home the point that persons cannot privilege or exempt themselves, their co-workers, or their favored group and still act morally.

Kant and contemporary Kantians have worked diligently on perhaps the single most important issue in recent moral philosophy: Are some actions wrong not because of their good or bad consequences, but because of the inherent wrongness of either the actions or the rules from which the action is performed? Also, Kant's second formulation of the categorical imperative—that persons must be treated as ends and not means only—can be, and often has been, interpreted as the substantive basis of the principle of respect for autonomy. Among the most defensible implications of his philosophy is that we have a basic, nonutilitarian obligation to respect the reasoned choices of others as well as their inherent capacities of reason and choice. Kant's formulation of this claim has deeply and justifiably influenced contemporary biomedical ethics.

RIGHTS THEORY

Utilitarian and Kantian theories are committed to the language of moral obligations, but the language of moral rights is no less important. Since at least the seventeenth century,[36] statements and theories of rights have been considered vital sources of the protection of life, liberty, expression, and property. They protect against oppression, unequal treatment, intolerance, insecurity, invasion of privacy, and the like. Many philosophers, political activists, lawyers, and framers of political declarations now regard rights theory as the most important type of moral theory.

An ethical analysis of the case of the five-year-old in need of a kidney transplant would, from the perspective of rights theory, focus on the rights of all parties in an effort to determine their meaning and scope as well as their weight and strength. The father could be considered to have rights of autonomy, privacy, and confidentiality that protect his bodily integrity and sphere of decision making from interference by others. In addition, he has a right to information, which he apparently received, about the risks, benefits, and alternatives of living kidney donation. The father's decision not to donate is within his rights, as long as it does not violate another's rights. No apparent grounds support a general right to assistance that could permit someone, including his daughter, to demand a kidney based on a moral right. However, there are various rights to assistance, and it could be argued that the daughter has a right to receive a kidney from her family on the basis of either parental obligations or medical need. If such a right exists, which is doubtful, it would be circumscribed. For example, it is implausible to suppose that such a right could be enforced against the interests of the girl's two young siblings. They are shielded from conscription of a kidney by their right of noninterference when the procedure is not for their direct benefit and carries risks.

The father has exercised his rights of autonomy and privacy in allowing the physician to run tests. He then seeks protection under his right of confidentiality, which he believes allows him to control third-party access to information generated in his relationship with the physician. However, the precise scope and limits of his rights and the competing rights of others need to be approached cautiously, as is the case with most presentations of abstract rights, especially in wide-ranging documents such as the United Nations' *Universal Declaration of Human Rights*.[37] For example, in the present case, does the mother herself have a right to the information generated in the relationship between the father and the nephrologist, particularly information bearing on the fate of her daughter? Another issue is whether the physician has a right of conscientious refusal. Even if the physician has a right to protect his personal integrity and to resist becoming an instrument of the father's desire to keep others from knowing why he is not donating, does this right trump the father's right of confidentiality?

Rights as Justified Claims

A right gives its holder a justified claim *to* something (an entitlement) and a justified claim *against* another party. Claiming is a mode of action that appeals to moral norms that allow persons to demand, affirm, or insist upon what is due to them. "Rights," then, are justified claims to something that individuals or groups can legitimately assert against other individuals, groups, or institutions. A right positions an individual or group to determine by one's choices what others morally must or must not do.[38]

The language of rights has served, on numerous occasions, as a means to oppose the status quo, to demand recognition and respect, and to promote social reforms that aim to secure legal protections for individuals. The

legitimate role of civil, political, and legal rights in protecting the individual from societal intrusions is undeniable, but the proposition that individual rights provide the fountainhead for moral and political theory has been resisted—for example, by many utilitarians and communitarians. They maintain that individual interests said to be protected by rights are often at odds with communal and institutional interests and also produce bizarre situations in which two or more rights claims are in direct conflict. In discussions of health care delivery, for example, proponents of a broad availability and distribution of medical services often appeal to the "right to health care," whereas opponents sometimes appeal to the "rights of the medical profession."

Many participants in moral, political, and legal discussions presuppose that arguments cannot be persuasive unless stated in the language of rights, a position we find unconvincing. Others find the language of rights excessively confrontational, adversarial, and unsuitable to address the moral problems that require attention. Although we also reject this position, we recognize that rights language is not well suited to handle all conflicts, and that not all claims to have a right have an adequate philosophical or legal basis.

Are Rights Trumps? Absolute and Prima Facie Rights

Rights are neither as strong nor as confrontational as many critics appear to claim. Some rights may be absolute or close to absolute[39] such as the moral right to consent to surgery or to choose one's religion or to reject all religion; but, typically, rights are not absolute claims. Like principles of obligation, rights assert only prima facie claims (in the sense of "prima facie," or, alternatively, "pro tanto"[40] as presented in Chapter 1, pp. 15–16).

Many writers in rights theory seem to dispute this claim. They regularly appeal to Ronald Dworkin's suggestive language that critical interests of individuals (chiefly when in conflict with the interests of political states) are firmly protected by rights that have the force of trump cards.[41] This trump metaphor is not well suited for situations in which one moral right conflicts with another moral right; and when individual rights conflict with the public interest these rights are not always trumps against the state. If the state needs to protect the rights of citizens—for example, the state needs to prevent the spread of a catastrophic disease—it may legitimately override some individual rights such as the right to refuse vaccination or to travel freely.

Dworkin gives a notably narrow account of rights as trumps: "Rights are best understood as *trumps over some background justification for political decisions* that states a goal for the community as a whole."[42] A right in this relatively narrow context is a claim persons can make regarding certain political actions or inactions. A trump right leaves one in a position to assert that governments cannot override rights for utilitarian or communitarian reasons, even if the action would maximally promote the public interest. In effect, Dworkin regards rights as stronger—much stronger—than the moral claims created by community goals and preferences that threaten these rights, especially as the claims have been developed in utilitarian theories. So understood, rights are instruments that function to guarantee that individuals cannot be sacrificed to government interests or mere majority interests, but they are not absolute trumps in any broader respect.

Interpreting rights as trumps is appealing in contexts in which individuals are vulnerable to serious harms and in which minority populations might be oppressed by majority preferences. The trump metaphor reminds us that rights powerfully protect individuals from having their interests balanced or traded off and that proposals to override them in the public interest need the most careful inspection and justification. However, models of rights based on trumps, absolute shields, and uninfringeable deontological protections can be more misleading than instructive and can be morally dangerous.

All rights, like all principles and rules of obligation, are prima facie, that is, presumptively valid claims that sometimes must yield to other claims. In light of this need to balance claims, we should distinguish a *violation* of a right from an *infringement* of a right.[43] "Violation" refers to an unjustified and wrong action against an interest that is protected by a right, whereas "infringement" refers to an action that may or may not legitimately override a right.

The Rights of Incompetent, Disadvantaged, and Unidentified Members of Populations

Possession of a right is independent of being in a position to assert the right or to exercise the right. A right-holder need not be the claimant in a particular case in order to have a justified claim. The fact that persons do not know that they have a right is no basis for asserting that they do not have it. Infants, the severely mentally handicapped, and oppressed ethnic and racial minority populations may not know, assert, or be able to claim their rights, but they still possess them, and claims can be made on their behalf by appropriate representatives. Likewise, many dependent humans and dependent laboratory animals may have rights whether or not they have an authorized representative, such as a surrogate, who can exercise the rights.

When problems about the rights of minority populations arise, it is sometimes difficult to tell whether the rights at stake are rights of individual members of the minority population or of the group itself—a distinction between group rights and individual rights that we encountered previously in the case of diabetes research on the Havasupai Indians (in Chapter 5, pp. 200–02).[44] In some circumstances obligations exist to protect rights even if no specific individuals and no specific group can be identified as being vulnerable to violations of their rights. For example, professionals in veterinary public health have obligations to protect both animals and people against communicable diseases, even though no specific animal or human is identifiable in many circumstances.[45] These right-holders are unidentified members of populations.

Similarly, it is sometimes difficult to impossible to determine which individuals have obligations to protect the welfare rights (rights to a decent minimum of well-being) of persons in situations of deprivation; and it is also difficult to determine which individuals have a right to relief from their circumstance of deprivation. Here persons lack claim-rights that can have any practical effect. These problems are immense in scope and profoundly real, especially in the situations of global injustice discussed in Chapter 7 (pp. 297–300). We have seen that when some party has a right, then someone else has a correlative obligation. But if the person with the obligation cannot be located by an applicable theory, law, or principle, it is doubtful under that theory, law, or principle that someone truly has a right.

These problems illustrate the importance of the specification of rights whenever specification can occur, usually through the offices of an institution with authority and resources.

Positive Rights and Negative Rights

The distinction between positive rights and negative rights has long been central in rights theory, but it achieved real prominence and salience in moral philosophy and in many nation-states only in the last quarter of the twentieth century. A major influence on both philosophical theory and politics was Henry Shue's 1980 book *Basic Rights: Subsistence, Affluence, and U.S. Foreign Policy*. Shue distinguishes between rights that protect individuals' security and "subsistence rights" such as the right to ample food and the right to adequate shelter. He argues that subsistence rights (which are positive rights) are as basic as security rights (which are negative rights). He also argues that no significant difference exists in their moral importance because both are basic rights.[46]

Shue's book is not a work in biomedical ethics, but it nicely illustrates what is at stake in the discussion of whether positive rights are as important and basic as negative rights in many areas of practical ethics. It also is a rich source of reflection on the distinction between positive obligations (to provide goods or services) and negative obligations (to refrain from harming). The value of these distinctions for both applied theory and public policy cannot be overstated, and, arguably, Shue's work has also made a major contribution to thinking about the place of basic rights in any adequate theory of global justice (see Chapter 7, pp. 297–300).

To move now to biomedical ethics, a *positive* right is a right to receive a particular good or service from others, for example, a right to health care and a right to public health protective services, whereas a *negative* right is a right to be free from some intervention by others, for example, a right of privacy and a right not to be forced to undergo involuntary institutionalization for psychiatric treatment. A person's positive right entails another's obligation to do something for that person; a negative right entails another's obligation to refrain from doing something.[47]

Some negative rights such as the right to refuse a recommended medical procedure or involvement in research are arguably grounded in the principle of respect for autonomy, whereas positive rights such as the right to health care are arguably grounded in principles of beneficence and justice. Although rights theorists have historically found it easier to justify negative rights, the modern recognition of welfare or entitlement rights has expanded the scope of positive rights in many nation-states and these rights are now widely discussed in analyses of justice in biomedical ethics. However, debates persist about which particular claims of positive and negative rights are morally justified and also about whether either is a basic human right.

The Correlativity of Rights and Obligations

How are rights related to the moral obligations that were prominently featured in the previous two moral theories treated in this chapter?

To answer this question, consider the meaning of the abstract statement "X has a right to do or have Y." Following from the earlier analysis of the nature of a right as a valid claim, X's right entails that some party has an obligation either not to interfere if X does Y or to provide X with Y. In all contexts of rights, a system of norms imposes an obligation either to act or to refrain from acting so that X can do or have Y. In this way, the language of rights is translatable into the language of obligations: A right entails an obligation, and an obligation entails a right. If, for example, a physician agrees to take John Doe as a patient and commences treatment, the physician incurs an obligation to Doe, and Doe gains a correlative right to treatment. Likewise, if a state has an obligation to provide goods such as food or health care to needy citizens, any citizen who meets the relevant criteria of need is entitled to an allotment of the food or health care. Here a strong connection exists between making demands and using rights to justify the demands. As Shue puts it, a "right provides the rational basis for a justified demand."[48]

The correlativity of obligations and rights is generally accepted in both philosophical ethics and legal theory, though the precise rights and obligations involved have been difficult to pin down. Here is a brief schema, using some basic rights and obligations, to illustrate the nature of correlativity:

Obligations	**Rights**
1. Do not kill.	1. The right to not be killed
2. Do not cause pain or suffering to others.	2. The right to not be caused pain or suffering by others
3. Prevent harm from occurring.	3. The right to have harms prevented from occurring
4. Rescue persons in danger.	4. The right to be rescued when in danger
5. Tell the truth.	5. The right to be told the truth
6. Nurture the young and dependent.	6. The right to be nurtured when young and dependent
7. Keep your promises.	7. The right to have promises kept
8. Do not steal.	8. The right to not have one's property stolen
9. Do not punish the innocent.	9. The right to not be punished when one is innocent
10. Obey the law.	10. The right to have others obey the law

Is the correlativity thesis flawed? The correlativity thesis has been challenged on grounds that the correlativity between obligations and rights is untidy[49] in that (1) only *some* obligations entail rights and (2) only *some* rights entail obligations.[50] We find these two challenges to correlativity unconvincing, but we will discuss only the first of the two challenges because many critics of the correlativity thesis now concede that all genuine rights (by contrast to merely proclaimed rights and aspirational rights) do carry correlative obligations. Also, the first challenge is the only crucial one for the theory that rights follow directly from obligations.

The objection is that various appropriate uses of the term *obligation*, as well as the related terms *requirement* and *duty*, show that some obligations do not imply correlative rights. Alleged examples come from the fact that we refer to obligations of charity, and yet no person can claim another person's charity as a matter of a right. Obligations of love and obligations of conscience are also put forward as examples of obligations without correlative rights.

The problem with these objections and counterexamples is that although it is correct to say that putative norms of "obligation" such as charity express what we "ought to do" or are "required to do" in some sense, these norms do not constitute genuine moral obligations. Rather, they obligate individuals who are committed to admirable moral ideals that exceed moral obligation. They are self-imposed rules of "obligation" that at bottom express widely admired and endorsed moral ideals rather than obligations imposed by morality. (See the discussion of moral ideals in Chapter 2, pp. 45–52.) The critical point is that all *genuine* (by contrast to *putative*) moral obligations do have correlative rights and that all genuine moral rights have correlative obligations.[51]

However, the line at which an action is obligatory rather than ideal is not always clear. Consider a circumstance in which a fire has broken out in a hospital. A child needs help to escape from a smoke-filled room. A physician sees the problem and takes the child from the room. The physician is not endangered in doing so and can easily carry the child to safety. Clearly this physician has a moral obligation to rescue the child—as would any passerby in the hall. However, if we alter the facts of this situation, it becomes questionable whether this moral obligation remains. Suppose the walls and floor of the hospital room are ablaze all around the child and collapse of the room will almost certainly occur at any second. The original obligation of beneficence now turns into a risky rescue mission that can be described as a moral "requirement" only in the misleading sense previously mentioned. In this circumstance, the physician has no obligation of rescue and the child has no right of rescue. As risks increase in circumstances of fires, epidemics, raging rivers, and other exceedingly dangerous circumstances, it becomes increasingly less likely that a genuine obligation exists, and a rescuer at some point in the risk index becomes a hero rather than a discharger of an obligation.

Are rights primary? The correlativity thesis does not determine whether rights or obligations, if either, is the more fundamental or primary category in moral theory. Proposals of a "rights-based" moral theory spring from a particular conception of the function and justification of morality.[52] If the objective of morality is to protect individuals' interests (rather than communal interests), and if rights (rather than obligations) are our primary instruments to this end, then moral action guides are fundamentally rights-based. In this philosophical account, rights precede and ground obligations.

A theory we encountered in Chapter 7 illustrates this position: Robert Nozick maintains that "individuals have rights, and there are things no person or group may do to them" without violating their rights.[53] He takes the following rule to be basic in the moral life: All persons have a right to be left free to do as they choose. The obligation not to interfere with this right follows from the right, rather than the right following from the obligation. That it follows in this way indicates the priority of a moral right over a rule of moral obligation; the obligation is entailed by the right.

Alan Gewirth has proposed a rights-based argument that recognizes *positive* or *benefit* rights that Nozick does not accept:

> Rights are to obligations as benefits are to burdens. For rights are justified claims to certain benefits, the support of certain interests of the subject or right-holder. Obligations, on the other hand, are justified burdens on the part of the respondent or duty-bearer; they restrict his freedom by requiring that he conduct himself in ways that directly benefit not himself but rather the right-holder. But burdens are for the sake of benefits, and not vice versa. Hence obligations, which are burdens, are for the sake of rights, whose objects are benefits. Rights, then, are prior to obligations in the order of justifying purpose … in that respondents have correlative obligations *because* subjects have certain rights.[54]

Such rights-based theories accept the correlativity of rights and obligations, but only when joined to a priority thesis that obligations follow from rights, rather than the converse. Rights form the justificatory basis of obligations because they best capture the objective of morality, which is to secure liberties or welfare benefits for a rights-holder.

Although we enthusiastically accept the correlativity thesis, we do not accept a priority thesis that renders rights, obligations, or virtues primary because we see no basis on which to sustain this philosophical conclusion.

The specification of rights. James Griffin rightly points out that we are sometimes satisfied that a basic right exists and that there are correlative obligations, yet we are uncertain precisely *what* the term "human rights" means and what a basic right gives us a right to.[55] Basic rights are abstract moral notions that do not fix how to formulate specific policies or resolve practical moral problems. We agree and also agree with Dworkin's assessment that "abstract rights … provide arguments for concrete rights, but the claim of a concrete right is more definitive [in political contexts] than any claim of abstract right that supports it."[56]

These problems should be handled through what we describe in several chapters as specification: the process of reducing the indeterminate character of abstract norms and giving them specific action-guiding content. Specifying rights to make them practical guidelines is just as important as specifying obligations.

A Critical Evaluation of Rights Theory

Problems in various rights theories will now be addressed.

Problems about the scope of morality. Pure rights-based accounts that aspire to be comprehensive moral theories run the risk of truncating our understanding of the richness of morality, because rights cannot account for the moral significance of motives, supererogatory actions, virtues, and the like. A moral theory premised exclusively on rights would fare poorly under the criteria of comprehensiveness and explanatory and justificatory power proposed at the beginning of this chapter. Accordingly, it is undesirable to limit the foundations of morality or moral theory to a rights-based model.

Problems about whether to exercise rights. Often a moral problem turns not on whether someone has a right, but on whether rights-holders should or should not *exercise* their rights. If a person says, "I know you have the right to do X, but you should not do it," this moral claim goes beyond a statement of a right. One's obligation or one's character, not one's right, is in question. This problem shows why rights theory needs to be buttressed by theories of obligation and virtue.

The neglect of communal goods. Rights theorists sometimes write as if social morality's major concern is to protect individual rights against governmental or other forms of communal intrusion. This vision is too limited for an ethical theory, however it may fare as a political theory. It excludes not only group interests but also communal values, such as public health, biomedical research, and the protection of animals used in research. The better perspective is that social ideals, principles of obligation, and communal interests are as central to morality as rights, and that none is dispensable.

A Constructive Evaluation of Rights Theory

We have offered a sympathetic interpretation of the use of rights language to express critically important, and universally valid, moral norms. We have also offered a defense of both the correlativity of rights and obligations and the moral and social purposes served by a theory of basic rights. No part of our moral vocabulary does more to protect the legitimate interests of citizens in political states than the language of rights. Predictably, injustice and inhumane treatment occur most frequently in political states that fail to recognize human rights in their rhetoric, documents, and actions. As much as any part of moral discourse, human rights language crosses national boundaries and enters into international law and statements by international agencies and associations.

Being a rights-bearer in a society that enforces rights is both a source of personal protection and a source of dignity and self-respect. By contrast, to maintain that someone has an obligation to protect another's interest may leave the beneficiary in a passive position, dependent on the other's goodwill in fulfilling the obligation. When persons possess enforceable rights correlative to obligations, they are enabled to be independent agents, pursuing their projects and making legitimate claims.

We value rights because, when enforced, they provide protections against unscrupulous behavior, promote orderly change and cohesiveness in communities, and allow diverse communities to coexist peacefully within a single political state.[57] A weighty reason for giving prominence to rights in moral and political theory is that in contexts of moral practice, such as health care institutions, rights demand a robust form of respect and better shield individuals against unjust or unwarranted communal intrusion, control, or neglect than does any other moral and legal category.

VIRTUE THEORY

In Chapter 2, we defended a theory of moral character that is delineated in terms of moral virtues. We now return to this domain to examine virtue theory as a major type of moral theory.

Virtue theory is independent of utilitarian, Kantian, and rights theories. Whatever their differences, utilitarians and deontologists similarly conceive of moral philosophy and the demands of morality: Ethics begins with the question, "What morally ought we to do?" and then provides general rules of obligation as guides to action. In classical Greek philosophy of the virtues, represented by Aristotle's powerful account, the cultivation of virtuous traits of character is conceived as one of morality's primary functions; and in the commanding eighteenth-century virtue theory of David Hume, even moral judgments of human actions are at bottom judgments of whether certain motives and character traits are virtuous or vicious.

Some defenders of virtue ethics deny that it is a theory at all, preferring to use terms such as *account* or *perspective* that can highlight virtue ethics' wide-ranging and all-embracing features. Others argue that placing virtue ethics together with the three theories thus far examined in this chapter loses sight of its radical critique of the other three approaches (and of contemporary culture).[58] Nevertheless, we regard virtue ethics as an alternative type of theory even if it does not address exactly the same questions as utilitarian, Kantian, or rights theories.[59]

We begin by considering how a proponent of virtue ethics might approach the case of the father who is reluctant to donate a kidney to his dying daughter and requests that the physician deceive his family about his reasons. The father's confessed lack of courage to donate one of his kidneys is relevant to an evaluation of him and his refusal to donate, but he had other reasons as well, some possibly involving self-deception. He points to his daughter's "degree of suffering," which suggests that he believes she might be better off without a transplant. Hence, his motives may be partially other-directed, not purely self-centered, and may involve compassion for his ill daughter. We could also investigate whether the father was compassionate and caring about her welfare and whether his apparent failure of courage may have overwhelmed his compassion, faithfulness, and other virtues, if they were present at all.

Several other judgments of character are relevant in assessing this case. We lack a full description of his wife, but the father was apparently worried that she would be unforgiving in accusing him of "allowing his daughter to die." This belief underlies his request that the physician tell a lie. In responding to the father's request, the physician focused on how the act of deception might compromise his integrity, and he "felt very uncomfortable" about the request. This feeling suggests an ongoing concern not to compromise his truthfulness and moral integrity. The physician presumably thought he could avoid a serious compromise of both truthfulness, in the sense of not directly lying, and integrity by saying that "for medical reasons" the father should not donate a kidney. However, questions arise about whether the physician acted on an unstable distinction between a direct lie (for instance, "he cannot donate because he is not histocompatible") and a deliberately misleading statement (for example, he should not donate "for medical reasons").

In the remainder of this section, we will first consider the critical distinction between right action and virtuous action, and then turn to the special status of virtues. We will later consider how moral virtues are related to the sorts of action guides presented in the previous three theories in this chapter.

Right Action and Proper Motive

Aristotle drew an important distinction between right action and proper motive, which he analyzed in terms of the distinction between external performance and internal state. An action can be right without being virtuous, he maintained, but an action can be virtuous only if performed in the right state of mind. Both right action and right motive are present in a truly virtuous action: "The agent must … be in the right state when he does [the actions]. First, he must know [that he is performing virtuous actions]; second, he must decide on them, and decide on them for themselves; and third, he must also do them from a firm and unchanging state," including the right state of emotion and desire. "The just and temperate person is not the one who [merely] does these actions, but the one who also does them in the way in which just or temperate people do them."[60]

Aristotle has it right. In addition to being properly motivated, a virtuous person experiences appropriate feelings, such as sympathy and regret, even when the feelings are not motives and no action results from the feelings. Virtuous persons also do not act from mere inclination or for personal advantage. They act under a conception of what is morally right and worthy. However, not all virtues have a transparent link to motives, feelings, or a conception of good and worthy reasons. Moral discernment and moral integrity, two virtues treated in Chapter 2, are examples. In these two virtues, psychological properties other than feelings are paramount, and they involve morally good states of mind in addition to having a conception of what is right and worthy.[61]

The terms *virtue* and *vice* are today less common in our everyday moral vocabulary than obligation, human rights, and the like, but the virtues have figured prominently in the history of both ethical theory and medical ethics. Appeals to virtue are intuitive and sensible: We commend and deeply respect persons who are honest, fair, respectful, just, or caring, or have various other admirable qualities. Likewise, we condemn and disrespect persons who are dishonest, malevolent, uncaring, unjust, or dishonorable, or have other vices. A comprehensive catalogue of the virtues and the vices, as proposed in some classic moral theories and religious traditions, is a large project, because there are dozens of vices and virtues.[62] Some are merely proclaimed virtues that are controversial, but many have been accepted in the common morality and by the major moral theorists who have developed accounts of virtue and vice.

The Definition of "Virtue"

The definition of "virtue" was briefly addressed in Chapter 2, where we stated that "A *virtue* is a dispositional trait of character that is socially valuable and reliably present in a person, and a *moral virtue* is a dispositional trait of character that is morally valuable and reliably present." This definition builds on, but moves beyond, a prominent definition of virtue offered by Hume, who wrote, "It is the nature, and, indeed, the definition of virtue, that it is *a quality of the mind agreeable to or approved of by every one, who considers or contemplates it.*"[63] So understood, a virtue is a fusion of two components: (1) an objective mental quality in a person (a feeling, motive, or character trait), and (2) a general approval of this mental quality by all impartial persons who contemplate it. "General approval" here means social approval of a type of mental trait such as benevolence, friendliness, gratitude, honesty, compassion, and public spiritedness. In Hume's theory, impartial moral judges (hence "every one" in his definition) are the sources of approval. A mental quality is a moral virtue if and only if it evokes the universal moral approval of impartial persons; and a mental quality is a vice if and only if the quality evokes universal condemnation in impartial persons. All morally decent persons have the capacity to see certain mental traits as estimable, agreeable, and amiable—to use Hume's terms.

Hume's definition provides the skeletal beginnings of an adequate analysis of "virtue." To generalize now beyond Hume's and Aristotle's immensely influential theories, a virtue is a deeply entrenched, morally good and socially commended trait of character that makes persons morally reliable, whereas a vice is the converse. We do not always think of virtue theory in terms of character traits, because parts of our vocabulary include "virtuous

action" and "virtuous person." Nonetheless, a moral virtue is itself a character trait. These traits dispose persons to perform right actions, but virtue theory proposes that we not start with right actions as if the virtues were derivative from judgments of action. The idea is to construct a moral theory from character traits that enable and dispose a person to identify and perform right actions.[64]

The Special Status of the Virtues

Some who write about virtue and character see the language of obligation as *derivative* from the language of virtue. They regard a person disposed by character to have good motives and desires as the model of the morally good person. This model determines our expectations of persons, which are then expressed in terms of their obligations.[65] They regard the virtue model as morally more basic and important than a model of action performed from obligation, because right motives and character tell us more about the moral worth of a person than do right actions goaded by obligation.

We are often more concerned about the character and motives of persons than about whether their acts conform to rules. When friends perform acts of "friendship," we expect the acts not to be motivated entirely from a sense of obligation to us, but to be motivated by a desire to be friendly accompanied by a sense of valuing our friendship. The friend who acts only from obligation lacks the virtue of friendliness, and in the absence of this virtue, the relationship lacks the moral quality of friendship.[66] Similar points hold for parents who play with their small children only because of a felt obligation.

Virtue theorists argue that the attempt in obligation-oriented theories to replace the virtuous judgments of health care professionals with rules, codes, and procedures—as has recently occurred in many professional codes—will not produce better decisions and actions.[67] For example, rather than relying on institutional rules and government regulations to protect human research subjects, the most reliable protection is the presence of an "informed, conscientious, compassionate, responsible researcher."[68] If so, character is more important and consequential than conformity to rules, and a premium should be placed on inculcating and cultivating the virtues through educational interactions and guidance by role models. Persons who are respectful, benevolent, and just are those who reliably perform right actions.

In his chronicle of life under the Nazi SS in the Jewish ghetto in Cracow, Poland, Thomas Keneally describes a physician faced with a moral dilemma: either inject cyanide into four immobile patients or abandon them to the SS, who were at that moment emptying the ghetto and had already demonstrated that they would torture and kill captives and patients. This physician, Keneally observes, "suffered painfully from a set of ethics as intimate to him as the organs of his own body."[69] Here is a person of the highest moral character and virtue, motivated to act rightly and even heroically, despite having no idea about the morally right action in this dilemmatic circumstance (given the lack of guidance found in traditional rules of medical ethics). Ultimately, with uncertainty and reluctance, the physician elected euthanasia, using forty drops of hydrocyanic acid, without the consent or knowledge of the four doomed patients—an act almost universally denounced by the canons of professional medical ethics. Even if one thinks that the physician's *actions* of killing were wrong and blameworthy—a judgment we would not defend—no reasonable person would make a judgment of blame or demerit directed at the physician's *motives* or *character.* Having already risked death by choosing to remain at his patients' beds in the hospital rather than take an available escape route, this physician is a moral hero who displayed an extraordinary moral character.

Action Guidance Based on Moral Virtues

What do virtuous moral agents do? Some virtue theorists maintain that virtues enable persons to discern what they should do and be motivated to do it in particular circumstances *without need for preexisting rules*. According to Rosalind Hursthouse:

> Virtue ethics provides a specification of "right actions"—as "what a virtuous agent would, characteristically, do in the circumstances"—and such a specification can be regarded as generating a number of moral rules or principles (contrary to the usual claim that virtue ethics does not come up with rules or principles). Each virtue generates an instruction—"Do what is honest," "Do what is charitable," and each vice a prohibition—"Do not … do what is dishonest, uncharitable."[70]

In this theory, what is right to do is what a virtuous agent would do, and the virtuous agent reliably does what conforms to a "virtue-rule." When moral conflicts and moral dilemmas of the sort explored in Chapter 1 emerge, they can be handled through additional specifications. Virtue ethics therefore resembles other normative ethical theories in seeking to identify the morally relevant features of a situation that justify doing action X rather than action Y.

Many proponents of virtue ethics do not lament that their approach lacks a clear and precise decision procedure for conflicts and dilemmas. They maintain that theories based on principles, rules, and rights have no advantage over virtue theory in resolving moral dilemmas; and they claim that, in irresolvable and tragic dilemmas, the virtues help direct agents to appropriate responses, including appropriate attitudes and emotions such as moral distress.[71]

Specification of the actual "instruction" or "virtue-rule" will often not be as straightforward as Hursthouse's examples may suggest (e.g., consider the virtue of moral integrity), and there is no reason to think that all specifications will rely exclusively on underlying notions of virtue. For example, rules of informed consent may rely on values of autonomy beyond the virtue of respectfulness for autonomy. Specification in virtue ethics is likely to be similar in its commitments to the theory of moral norms and specification that we proposed in Chapter 1. Virtue theory, from this perspective, does not prove that virtues have advantages over principles and rules of obligation as guides to action.

The moral life is a constant process of acquiring skills and making judgments. Over time a person gains increased understanding and becomes more skilled in specifying general guidelines, becoming morally virtuous, and sticking to one's moral ideals. In the case of virtues and moral ideals, one learns better how to be truthful, honest, discreet, friendly, charitable, and polite by bringing those virtues to bear in a variety of situations. This way of learning involves the acquisition of skills roughly analogous to the process of learning how to use a language.[72]

The correspondence of moral virtues and moral obligations. There is a rough, although imperfect, correspondence between some virtues and moral principles, rules, and ideals. This relationship is less uniform and more complicated than the correlativity of rights and obligations discussed in the previous section of this chapter. The following (noncomprehensive) list illustrates the correspondence between a few select virtues and norms that are prominent in our account of the common morality.

Principles	**Virtues**
Respect for autonomy	Respectfulness for autonomy
Nonmaleficence	Nonmalevolence
Beneficence	Benevolence
Justice	Justice
Rules	**Virtues**
Veracity	Truthfulness
Confidentiality	Respectfulness for confidentiality
Privacy	Respectfulness for privacy
Fidelity	Faithfulness
Ideals of Action	**Ideals of Virtue**
Exceptional forgiveness	Exceptional forgivingness

Exceptional generosity	Exceptional generousness
Exceptional compassion	Exceptional compassionateness
Exceptional kindness	Exceptional kindliness

This list could be expanded to include an extensive array of additional norms and virtues, but no table can be constructed that presents a perfect, comprehensive schema of correspondence and noncorrespondence. Moreover, many virtues do not have a direct, one-to-one correspondence to a principle. For example, caring, concern, compassion, sympathy, courage, modesty, and patience are virtues that do not correspond well to principles and rules of obligation. Other examples are cautiousness, integrity, cheerfulness, unpretentiousness, sincerity, appreciativeness, cooperativeness, and commitment.[73] Some virtues that lack corresponding norms of obligation nonetheless have corresponding moral ideals, as the above list indicates. And all are important for morality as a whole.

A Critical Evaluation of Virtue Theory

Several problems merit consideration in assessing virtue theory.

How independent and comprehensive is virtue theory? Various virtues seem to be character traits compatible with the performance of morally *wrong* actions. For example, courage, wisdom, and loyalty can enable unethical activities. As discussed in Chapter 8, the virtues of loyalty, friendship, and solidarity can foster inadequate reporting by physicians of unethical or incompetent behavior by other physicians. In speaking of generally admirable character traits as *moral* virtues, virtue theory cannot merely list good, commendable, and useful mental traits.

In the tradition descending from Aristotle, a moral virtue is exclusively a moral excellence of a person. But is moral excellence or moral worthiness determined exclusively by virtue standards? The notion of a morally worthy pursuit is not analyzable entirely in terms of a virtue theory because it will often rely on nonvirtue premises regarding what constitutes a morally good life and morally good conduct, which in turn may require reference to action guides and the basic objectives of morality.

When strangers meet. Virtue and character are likely to be prized and emphasized in many human relationships where a climate of trust prevails. Principles or rules that express the obligations of health professionals in codes of conduct and statements of patients' rights may, in these intimate contexts, be intrusions rather than essential elements. However, virtue theory works less well for certain forms of moral encounter, especially where trust, intimacy, familiarity, and the like have not been established. When strangers meet, character often plays a less significant role than principles, rules, and institutional policies. For example, when a patient first encounters a physician, the physician's conformity to rules may be essential in situations of obtaining consent, disclosing a conflict of interest, proposing "do not resuscitate" orders for incompetent patients, explaining surrogate mother arrangements, and so on. Likewise, physicians may welcome explicit and mutually agreed-upon rules of informed consent, advance directives, codes of ethics, and similar structures and arrangements. Here rights, rules, and guidelines are welcome and are entirely acceptable parts of the moral landscape.

A Constructive Evaluation of Virtue Theory

Virtues come to the fore in contexts in which trust, intimacy, and dependence are present. Virtue theory is particularly well suited to help us navigate circumstances of caregiving and the delivery of information in health care. For example, "consenting a patient" (a common but problematic expression)[74] by merely conforming to institutional rules of informed consent is generally far less important than having a caring and discerning physician, nurse, or other health professional who appreciates the importance of dialogue, reassurance, and honesty in the process of obtaining an informed consent.

Virtue theory is the most venerable type of moral theory, with a grand tradition descending from the ancient to the modern world. Throughout the history of moral theory, leading writers on the virtues have agreed on most of

the moral virtues and on the importance of virtue theory. Aristotle's emphasis on excellences of character and David Hume's emphasis on virtues as the basis of personal moral merit are jewels in the history of virtue theory and moral philosophy. Though 2,000 years separated them, their philosophies display a considerable agreement about the central virtues as well as about the centrality of virtue theory in moral philosophy. These theories deserve a status and recognition no less prominent than Mill's utilitarian views of social beneficence, Kant's deontological views about the categorical demands of respect for all persons, and the views of celebrated writers in the history of rights theory.

THE CONVERGENCE OF THEORIES ON PRINCIPLES

When competing theories, systems, or general depictions of some phenomenon are available, we usually seek out the best account. However, affiliation with a single type of ethical theory is precarious, especially in biomedical ethics. If the two authors of this book were forced to rank the types of theory examined in this chapter, we would differ. Nevertheless, for both of us, the most satisfactory type of theory—if we could find *one* to be most satisfactory—would be only slightly preferable, and no theory would fully satisfy all of the criteria for assessing theories presented in the first section of this chapter.

Differences among types of theory should not be exaggerated because these theories are not analogous to warring armies locked in combat. Many and perhaps most moral theories lead to the acceptance of the same general action guides that we have presented in several chapters as elements in the common morality. This thesis may work less well for act-based theories (notably for act utilitarianism), but it generally holds for theories committed to principles, rules, rights, and virtues. These different theories often defend roughly the same principles, obligations, rights, responsibilities, virtues, and the like. For example, rule utilitarianism may appear to be starkly different from, and even hostile to, nonconsequentialist theories, but rule-utilitarian Richard Brandt rightly notes that his theory is similar, at the level of principle and obligation, to W. D. Ross's nonutilitarian theory (discussed in Chapter 1):

> [The best code] would contain rules giving directions for recurrent situations which involve conflicts of human interests. Presumably, then, it would contain rules rather similar to W. D. Ross's list of prima facie obligations: rules about the keeping of promises and contracts, rules about debts of gratitude such as we may owe to our parents, and, of course, rules about not injuring other persons and about promoting the welfare of others where this does not work a comparable hardship on us.[75]

That Brandt appeals to utility and Ross to deontological considerations to justify similar sets of rules is a significant difference at the level of moral theory and justification. Brandt and Ross also might interpret, specify, and balance their rules differently as a result of their theoretical commitments, but their lists of primary obligations display no major differences. Their convergence on general principles is not unusual in moral theory. Agreement derives from an initial shared database, namely, the norms of the common morality. The proponents of the different types of theory examined in this chapter all accept the principles of common morality *before* they devise their theory—as we believe is also true of Aristotle, Locke, Hume, Kant, Mill, and other giants in the history of moral philosophy we have mentioned. This claim does not ignore the important differences that may appear in their interpretation and weighting of those principles.

Convergence on a basic set of norms is also common in assessing cases and framing policies, even if theoretical differences divide the discussants. As bioethics commissions and committees have long appreciated, in making practical judgments and creating public policies, we need only agreement on a set of basic action-guides—not an agreement on either their theoretical foundations or where they should and should not be applied. Nonetheless, convergence to agreement on general norms should not be confused with questions about whether a theory adequately justifies its principles. Theoretical inquiry is worthwhile even if practical agreement and serious moral progress in biomedical ethics can often be achieved without resolving deep theoretical differences.

CONCLUSION

Competition exists among the four types of normative theory explored in this chapter, and conflicting conceptions continue regarding what these theories imply for biomedical practice. However, each of these theories is instructive and makes a contribution to our thinking about the moral life. We have maintained that no reason exists to consider one type of theory inferior to or derivative from another, and there is good reason to believe that these types of theory all show considerable insight into our common moral heritage and how it can be mined to help us develop contemporary biomedical ethics.

Every general theory risks clashing at some point with considered moral convictions, but each of the four theories examined in this chapter articulates a point of view that we should be reluctant to relinquish. This approach to theories allows us to focus on their remarkable insights without being forced to choose one theory to the exclusion of the others or to judge one theory as primary at the foundations of ethics.

NOTES

1. 1. Our views on pluralism are influenced by Thomas Nagel, "The Fragmentation of Value," in *Mortal Questions* (Cambridge: Cambridge University Press, 1979), pp. 128–37; and Baruch Brody's treatment in *Life and Death Decision Making* (New York: Oxford University Press, 1988), especially p. 9.
2. 2. Our presentation has profited from Shelly Kagan, *The Limits of Morality* (Oxford: Clarendon Press, 1989), esp. pp. 11–15, and from criticisms of our views privately presented by David DeGrazia and Avi Craimer.
3. 3. For analysis of this utilitarian thesis, see Samuel Scheffler, *Consequentialism and Its Critics* (Oxford: Clarendon Press, 1988).
4. 4. Jeremy Bentham, *An Introduction to the Principles of Morals and Legislation*, ed. J. H. Burns and H. L. A. Hart (Oxford: Clarendon Press, 1970), pp. 11–14, 31, 34; and John Stuart Mill, *Utilitarianism*, in vol. 10 of the *Collected Works of John Stuart Mill* (Toronto: University of Toronto Press, 1969), chap. 1, p. 207; chap. 2, pp. 210, 214; chap. 4, pp. 234–35.
5. 5. See a representative theory in James Griffin, *Well-Being: Its Meaning, Measurement and Moral Importance* (Oxford: Clarendon, 1986), especially p. 67. The most influential early twentieth-century theory of this sort was G. E. Moore, *Principia Ethica*; see the revised edition, ed. Thomas Baldwin (Cambridge: Cambridge University Press, 1993).
6. 6. This case is based on Melvin D. Levine, Lee Scott, and William J. Curran, "Ethics Rounds in a Children's Medical Center: Evaluation of a Hospital-Based Program for Continuing Education in Medical Ethics," *Pediatrics* 60 (August 1977): 205.
7. 7. Influential utilitarian works in bioethics include Peter Singer, *Practical Ethics*, 2nd ed. (Cambridge: Cambridge University Press, 1993); R. M. Hare, *Moral Thinking: Its Levels, Method, and Point* (Oxford: Oxford University Press, 1981); Hare, *Essays on Bioethics* (Oxford: Oxford University Press, 1993); Hare, "A Utilitarian Approach to Ethics," in *A Companion to Bioethics*, ed. Helga Kuhse and Peter Singer, 2nd ed. (Oxford: Wiley-Blackwell, 2009), pp. 85–90; and Brad Hooker, *Ideal Code, Real World: A Rule-Consequentialist Theory of Morality* (Oxford: Oxford University Press, 2002). John Harris's influential work tends in a consequentialist and utilitarian direction; see Harris, *The Value of Life: An Introduction to Medical Ethics* (New York: Routledge, 1985), among other publications. Jonathan Baron, under the influence of Hare (among others) and decision theory, argues for a utilitarian approach to bioethics and against principlism. See Baron, *Against Bioethics* (Cambridge, MA: MIT Press, 2006).
8. 8. Cf. L. W. Sumner, *The Moral Foundation of Rights* (Oxford: Clarendon Press, 1987); and Hooker, *Ideal Code, Real World.*
9. 9. Worthington Hooker, *Physician and Patient* (New York: Baker & Scribner, 1849), pp. 357ff, 375–81.
10. 10. J. J. C. Smart, *An Outline of a System of Utilitarian Ethics* (Melbourne: Melbourne University Press, 1961); and Smart, "Extreme and Restricted Utilitarianism," in *Contemporary Utilitarianism*, ed. Michael D. Bayles (Garden City, NY: Doubleday, 1968), esp. pp. 104–7, 113–15.
11. 11. Richard B. Brandt, "Toward a Credible Form of Utilitarianism," in *Contemporary Utilitarianism*, ed. Bayles, pp. 143–86; and Brandt's *Morality, Utilitarianism, and Rights* (Cambridge: Cambridge University Press, 1992). For a rule-utilitarian alternative to Brandt's rule-utilitarian formulations, see Hooker, *Ideal World, Real World.*

12. 12. For wide-reaching analyses of utilitarianism, including critical assessments, see Tim Mulgan, *Understanding Utilitarianism* (Abingdon, UK: Routledge, 2014); and Walter Sinnott-Armstrong, "Consequentialism," *The Stanford Encyclopedia of Philosophy* (Winter 2015 Edition), ed. Edward N. Zalta, available at https://plato.stanford.edu/archives/win2015/entries/consequentialism/ (accessed April 8, 2018).
13. 13. This question is discussed in Madison Powers, "Repugnant Desires and the Two-Tier Conception of Utility," *Utilitas* 6 (1994): 171–76.
14. 14. Alan Donagan, "Is There a Credible Form of Utilitarianism?" in *Contemporary Utilitarianism*, ed. Bayles, pp. 187–202. See also the attempt to develop a consequentialist theory that reduces or eliminates the "demandingness" problem in Tim Mulgan, *The Demands of Consequentialism* (Oxford: Clarendon Press, 2005), which offers a "moderately demanding" theory of mixed consequentialism.
15. 15. Williams, "A Critique of Utilitarianism," in *Utilitarianism: For and Against*, ed. J. J. C. Smart and Bernard Williams (Cambridge: Cambridge University Press, 1973), pp. 116–17; and J. L. Mackie, *Ethics: Inventing Right and Wrong* (New York: Penguin, 1977), pp. 129, 133. For an extension, see Edward Harcourt, "Integrity, Practical Deliberation and Utilitarianism," *Philosophical Quarterly* 48 (1998): 189–98.
16. 16. For a defense of utilitarianism (set against egalitarianism) in forming just policies toward people with disabilities, see Mark S. Stein, *Distributive Justice and Disability: Utilitarianism against Egalitarianism* (New Haven, CT: Yale University Press, 2006).
17. 17. Milton C. Weinstein and William B. Stason, *Hypertension* (Cambridge, MA: Harvard University Press, 1977); "Public Health Rounds at the Harvard School of Public Health: Allocation of Resources to Manage Hypertension," *New England Journal of Medicine* 296 (1977): 732–39; and "Allocating Resources: The Case of Hypertension," *Hastings Center Report* 7 (October 1977): 24–29.
18. 18. We agree with Amartya Sen that "consequentialist reasoning may be fruitfully used even when consequentialism as such is not accepted. To ignore consequences is to leave an ethical story half told." *On Ethics and Economics* (Oxford: Basil Blackwell, 1987), p. 75.
19. 19. See Stephen Darwall, ed., *Deontology* (Oxford: Blackwell, 2003), for a representative collection of works; and Larry Alexander and Michael Moore, "Deontological Ethics," *The Stanford Encyclopedia of Philosophy* (Winter 2016 Edition), ed. Edward N. Zalta, available at https://plato.stanford.edu/archives/win2016/entries/ethics-deontological/ (accessed April 8, 2018).
20. 20. See, for example, the writings of F. M. Kamm, especially *Intricate Ethics: Rights, Responsibilities, and Permissible Harm* (New York: Oxford University Press, 2007). Rather than taking a specifically Kantian approach, she indicates that contemporary nonconsequentialism has "its spiritual roots in the work of Immanuel Kant and W. D. Ross" (p. 10). Her rigorous writings often focus on issues in or important for bioethics, as in her *Bioethical Prescriptions: To Create, End, Choose, and Improve Lives* (New York: Oxford University Press, 2013). Some nonconsequentialists or deontologists operate from Jewish, Christian, Islamic, or other religious perspectives, which we do not examine in this book.
21. 21. Kant sought to show that what we should do morally is determined by what we would do "if reason completely determined the will." *The Critique of Practical Reason*, trans. Lewis White Beck (New York: Macmillan, 1985), pp. 18–19; Ak. 20. "Ak." designates the page-reference system of the twenty-two-volume Preussische Akademie edition conventionally cited in Kant scholarship.
22. 22. Kant, *Foundations of the Metaphysics of Morals*, trans. Lewis White Beck (Indianapolis, IN: Bobbs-Merrill, 1959), pp. 37–42; Ak. 421–24.
23. 23. For interpretations of Kant's idea of contradiction in maxims, see Christine Korsgaard, "Kant's Formula of Universal Law," *Pacific Philosophical Quarterly* 66 (1985): 24–47, and "Kant's Formula of Humanity," *Kant-Studien* 77 (1986): 183–202, both reprinted with other essays in her *Creating the Kingdom of Ends* (Cambridge: Cambridge University Press, 1996); and Barbara Herman, *The Practice of Moral Judgment* (Cambridge, MA: Harvard University Press, 1993), pp. 132–58.
24. 24. Kant, *Foundations*, p. 47; Ak. 429.
25. 25. Kant, *Foundations*, pp. 51, 58–63; Ak. 432, 439–44.
26. 26. Kant, *Foundations*, p. 58; Ak. 439–40.
27. 27. Alan Donagan, *The Theory of Morality* (Chicago: University of Chicago Press, 1977), pp. 63–66.
28. 28. See *A Theory of Justice* (Cambridge, MA: Harvard University Press, 1971; rev. ed., 1999), pp. 3–4, 27–31 (1999: pp. 3–4, 24–28). For an approach to Kant influenced by Rawls, see Thomas Hill, Jr., *Human*

Welfare and Moral Worth: Kantian Perspectives (Oxford: Clarendon, 2002).

29. 29. Rawls, *A Theory of Justice*, pp. 252, 256, 515–20 (1999 ed.: pp. 221–22, 226–27, 452–56). See also his "A Kantian Conception of Equality," *Cambridge Review* (February 1975): 97ff.
30. 30. See, for example, Thomas Nagel, "Personal Rights and Public Space," *Philosophy & Public Affairs* 24 (1995): 83–107, and his *The View from Nowhere* (New York: Oxford University Press, 1986); Bernard Williams, *Ethics and the Limits of Philosophy* (Cambridge, MA: Harvard University Press, 1985), and his *Moral Luck: Philosophical Papers, 1973–1980* (Cambridge: Cambridge University Press, 1981).
31. 31. Christine M. Korsgaard, "Interacting with Animals: A Kantian Account," in *Oxford Handbook of Animal Ethics*, ed. Tom L. Beauchamp and R. G. Frey (New York: Oxford University Press, 2011), p. 97.
32. 32. Onora O'Neill, *Towards Justice and Virtue: A Constructive Account of Practical Reasoning* (Cambridge: Cambridge University Press, 1996), pp. 5–6; and *Constructions of Reason: Explorations of Kant's Practical Philosophy* (Cambridge: Cambridge University Press, 1989). Her Kantian work in bioethics includes *Autonomy and Trust in Bioethics* (Cambridge: Cambridge University Press, 2002), and, with Neil C. Manson, *Rethinking Informed Consent in Bioethics* (Cambridge: Cambridge University Press, 2007).
33. 33. For innovative interpretations that respond to this objection by giving more flexibility to Kant, see Herman, *The Practice of Moral Judgment*, pp. 132–58; Nancy Sherman, *Making a Necessity of Virtue* (Cambridge: Cambridge University Press, 1997); and Tamar Schapiro, "Kantian Rigorism and Mitigating Circumstances," *Ethics* 117 (2006): 32–57. These writings respond to forms of the third objection we mention in this section, especially regarding the place of virtue in Kant's theory.
34. 34. Cf. Annette Baier, "The Need for More than Justice," in her *Moral Prejudices* (Cambridge, MA: Harvard University Press, 1994).
35. 35. We are indebted to the analysis in Karen Stohr, "Virtue Ethics and Kant's Cold-Hearted Benefactor," *Journal of Value Inquiry* 36 (2002): 187–204.
36. 36. Pioneering theories of international rights and natural rights—now often restyled as *human* rights—first prospered in philosophy through the social and political theories of Hugo Grotius, Thomas Hobbes, John Locke, and their near successors, which were often contractarian theories. On the history of human rights, see Anthony Pagden, "Human Rights, Natural Rights, and Europe's Imperial Legacy," *Political Theory* 31 (2003): 171–99; the wide-ranging historical, anthropological, and philosophical theory in Ian Shapiro, *The Evolution of Rights in Liberal Theory* (Cambridge: Cambridge University Press, as reissued in 2008); and James Nickel, "Human Rights," *The Stanford Encyclopedia of Philosophy* (Spring 2017 Edition), ed. Edward N. Zalta, available at https://plato.stanford.edu/archives/spr2017/entries/rights-human (accessed April 10, 2018). For an exploration of human rights in relation to global bioethical issues, see Wanda Teays, John-Stewart Gordon, and Alison Dundes Renteln, eds., *Global Bioethics and Human Rights: Contemporary Issues* (Lanham, MD: Rowman & Littlefield, 2014). For an account of human rights closely connected to theories of justice and sometimes to bioethics, see Madison Powers and Ruth R. Faden, *Structural Injustice: Power, Advantage, and Human Rights* (New York: Oxford University Press, 2019). For a critique of appeals to human rights in bioethics, see two articles by John D. Arras and Elizabeth M. Fenton: "Bioethics and Human Rights: Access to Health-Related Goods," *Hastings Center Report* 39 (2009): 27–38; and "Bioethics and Human Rights: Curb Your Enthusiasm," *Cambridge Quarterly of Healthcare Ethics* 19 (2010): 127–33.
37. 37. United Nations, *Universal Declaration of Human Rights*, 2015 online edition, available at http://www.un.org/en/udhrbook/pdf/udhr_booklet_en_web.pd (accessed July 29, 2018).
38. 38. Our representations on this point are indebted to the theory of rights in Joel Feinberg's *Rights, Justice, and the Bounds of Liberty* (Princeton, NJ: Princeton University Press, 1980), esp. pp. 139–41, 143–55, 159–60, 187; and Feinberg, *Social Philosophy* (Englewood Cliffs, NJ: Prentice-Hall, 1973), chaps. 4–6. See also Alan Gewirth, *The Community of Rights* (Chicago: University of Chicago Press, 1996), pp. 8–9; H. L. A. Hart, "Bentham on Legal Rights," in *Oxford Essays in Jurisprudence*, 2nd series, ed. A. W. B. Simpson (Oxford: Oxford University Press, 1973), pp. 171–98; and Christian Reus-Smit, "On Rights and Institutions," in *Global Basic Rights*, ed. Charles Beitz and Robert E. Goodin (New York: Oxford University Press, 2009), esp. pp. 27–29.
39. 39. A clever and atypical attempt to defend an absolute right is Alan Gewirth, "Are There Any Absolute Rights?" *Philosophical Quarterly* 31 (1981): 1–16; reprinted in Gewirth's *Human Rights* (Chicago: University of Chicago Press, 1982), Chapter 9.

40. 40. On the distinction between absolute and prima facie rights, see Danny Frederick, "Pro-Tanto versus Absolute Rights," *Philosophical Forum* 45 (2014): 375–94.
41. 41. Ronald Dworkin, *Taking Rights Seriously* (Cambridge, MA: Harvard University Press, 1977), pp. xi, xv, 92 (and, as reissued with an "Appendix: A Reply to Critics," in 2002, pp. 364–66); and *Law's Empire* (Cambridge, MA: Harvard University Press, 1986), p. 160.
42. 42. Ronald Dworkin, "Rights as Trumps," in *Theories of Rights*, ed. Jeremy Waldron (Oxford: Oxford University Press, 1984), pp. 153–67; the quote is on p. 153 (italics added).
43. 43. See Judith Jarvis Thomson, *The Realm of Rights* (Cambridge, MA: Harvard University Press, 1990), pp. 122–24, and also 106–17, 149–53, 164–75; and Feinberg, *Rights, Justice, and the Bounds of Liberty*, pp. 229–32.
44. 44. On problems of minority rights, see James Nickel, *Making Sense of Human Rights*, 2nd ed. (Malden, MA: Blackwell, 2007), chap. 10. For the case for group rights, see James Griffin, *On Human Rights* (Oxford: Oxford University Press, 2008), chap. 15.
45. 45. World Health Organization, "Zoonoses: Managing Public Health Risks at the Human-Animal-Environment Interface," available at http://www.who.int/zoonoses/en/ (accessed April 10, 2018).
46. 46. Shue's first edition was published in 1980, and a second edition in 1996 (Princeton, NJ: Princeton University Press). The second is used here. The first edition was highly influential in several disciplines. The nature and importance of Shue's work is carefully explored by several authors in *Global Basic Rights*, ed. Beitz and Goodin.
47. 47. See Feinberg, *Social Philosophy*, p. 59; Eric Mack, ed., *Positive and Negative Duties* (New Orleans, LA: Tulane University Press, 1985); and Judith Lichtenberg, "Are There Any Basic Rights," in *Global Basic Rights*, ed. Beitz and Goodin, esp. pp. 81–91.
48. 48. Shue, *Basic Rights*, p. 13.
49. 49. See David Braybrooke, "The Firm but Untidy Correlativity of Rights and Obligations," *Canadian Journal of Philosophy* 1 (1972): 351–63; Feinberg, *Rights, Justice, and the Bounds of Liberty*, pp. 135–39, 143–44; Feinberg, *Harm to Others*, vol. 1 of *The Moral Limits of the Criminal Law* (New York: Oxford University Press, 1984), pp. 148–49; Griffin, *On Human Rights*, pp. 51, 96, 107–9; and Joseph Raz, *The Morality of Freedom* (New York: Oxford University Press, 1986), pp. 170–72. Probing discussions of correlativity are found in Gewirth's *The Community of Rights.* See also Feinberg's insightful explanation of the confusions that enter moral discourse owing to the ambiguity of the words *duty*, *obligation*, and *requirement*, in his *Doing and Deserving: Essays in the Theory of Responsibility* (Princeton, NJ: Princeton University Press, 1970), pp. 3–8.
50. 50. See some objections presented by David Lyons, "The Correlativity of Rights and Duties," *Nous* 4 (1970): 45–55; Theodore M. Benditt, *Rights* (Totowa, NJ: Rowman & Littlefield, 1982), pp. 6–7, 23–25, 77; Alan R. White, *Rights* (Oxford: Clarendon Press, 1984), pp. 60–66; and Richard Brandt, *Ethical Theory* (Englewood Cliffs, NJ: Prentice Hall, 1959), pp. 439–40.
51. 51. This difference is sometimes marked by saying that perfect obligations have correlative rights, whereas imperfect obligations do not. We prefer the tidier approach that only perfect obligations are genuinely moral obligations. So-called imperfect obligations are moral ideals that allow for discretion. Cf. Feinberg, *Rights, Justice, and the Bounds of Liberty*, pp. 138–39, 143–44, 148–49. For an interesting argument for a right to be loved, see S. Matthew Liao, *The Right to Be Loved* (New York: Oxford University Press, 2015).
52. 52. Ronald Dworkin argues that political morality is rights-based in *Taking Rights Seriously*, pp. 169–77, esp. p. 171. J. L. Mackie's theory develops a similar thesis applied to morality in general in "Can There Be a Right-Based Moral Theory?" *Midwest Studies in Philosophy* 3 (1978), esp. p. 350.
53. 53. Robert Nozick, *Anarchy, State, and Utopia* (New York: Basic Books, 1974), pp. ix, 149–82.
54. 54. Alan Gewirth, "Why Rights Are Indispensable," *Mind* 95 (1986): 329–44, quote from p. 333. See Gewirth's later book, *The Community of Rights* (Chicago: University of Chicago Press, 1996).
55. 55. James Griffin, *On Human Rights* (Oxford: Oxford University Press, 2008), pp. 14–19, 97, 110. Griffin maintains that "the term 'human right' is nearly 'criterionless'" (p. 14). See, for a different view, Joseph Raz, *The Morality of Freedom* (Oxford: Clarendon Press, 1986), chap. 7.1.
56. 56. Ronald Dworkin, *Taking Rights Seriously*, pp. 93–94.
57. 57. See William R. Lund, "Politics, Virtue, and the Right to Do Wrong: Assessing the Communitarian Critique of Rights," *Journal of Social Philosophy* 28 (1997): 101–22; Allen Buchanan, "Assessing the

Communitarian Critique of Liberalism," *Ethics* 99 (July 1989): 852–82, esp. 862–65; and William A. Galston, *Liberal Purposes* (Cambridge: Cambridge University Press, 1991).

58. 58. See Talbot Brewer, *The Retrieval of Ethics* (Oxford: Oxford University Press, 2009), pp. 1–11, passim.
59. 59. For an introduction, see Heather Battaly, *Virtue* (Cambridge: Polity Press, 2015). Several volumes provide a range of illuminating views: see, for example, Lorraine Besser-Jones and Michael Slote, eds., *The Routledge Companion to Virtue Ethics* (London: Routledge, 2015); and Daniel C. Russell, ed., *The Cambridge Companion to Virtue Ethics* (Cambridge: Cambridge University Press, 2013).
60. 60. Aristotle, *Nicomachean Ethics*, trans. Terence Irwin (Indianapolis, IN: Hackett, 1985), $1105^{a}17–33$, $1106^{b}21–23$; cf. also $1144^{a}14–20$.
61. 61. Robert Adams distinguishes "motivational virtues" (such as benevolence) from "structural virtues" (such as courage and self-control). The latter are structural features of the agent's organization and management of his or her motives. *A Theory of Virtue: Excellence in Being for the Good* (Oxford: Clarendon Press, 2006), pp. 33–34, passim.
62. 62. Categorization has centuries of tradition behind it in ethical theory. Although there are variations in the proposed lists of virtues (and vices), much is also held in common across these traditions—enough to speak of a common morality of the virtues. See David Hume's comments on the catalogue of the virtues in his *An Enquiry concerning the Principles of Morals*, ed. Tom L. Beauchamp (Oxford: Clarendon Press, 1998), beginning at 1.10 (sect. 1, par. 10); see also 6.21, 9.3, 9.12. Although Hume was influenced by Aristotle, he reported that he was most deeply influenced by the catalogue of virtues in Cicero's *De officiis.* For a description of positive character traits interpreted as virtues, see Christopher Peterson and Martin E. P. Seligman, eds., *Character Strengths and Virtues: A Handbook and Classification* (Washington, DC: American Psychological Association; and New York: Oxford University Press, 2004). Their chapters identify twenty-four specific character strengths under six broad virtues.
63. 63. Hume, *An Enquiry concerning the Principles of Morals*, sect. 8, footnote to the section title; and appendix 1, par. 10.
64. 64. For further analysis of the nature and definition of virtue, see Julia Annas, *Intelligent Virtue* (New York: Oxford University Press, 2011), esp. chaps. 2–5.
65. 65. See Philippa Foot, *Virtues and Vices* (Oxford: Basil Blackwell, 1978); Gregory Trianosky, "Supererogation, Wrongdoing, and Vice," *Journal of Philosophy* 83 (1986): 26–40; Jorge L. Garcia, "The Primacy of the Virtuous," *Philosophia* 20 (1990): 69–91; and criticisms of this perspective in Lynn A. Jansen, "The Virtues in Their Place: Virtue Ethics in Medicine," *Theoretical Medicine* 21 (2000): 261–76.
66. 66. See Diane Jeske, "Friendship, Virtue, and Impartiality," *Philosophy and Phenomenological Research* 57 (1997): 51–72; and Michael Stocker, "The Schizophrenia of Modern Ethical Theories," *Journal of Philosophy* 73 (1976): 453–66. On the history and central role of friendship in virtue theory dating from Aristotle, see Bennett Helm, "Friendship," *The Stanford Encyclopedia of Philosophy* (Fall 2017 Edition), ed. Edward N. Zalta, available at https://plato.stanford.edu/archives/fall2017/entries/friendship/ (accessed April 4, 2018); and Sandra Lynch, *Philosophy and Friendship* (Edinburgh: Edinburgh University Press, 2005).
67. 67. Cf. Gregory Pence, *Ethical Options in Medicine* (Oradell, NJ: Medical Economics, 1980), p. 177.
68. 68. The quotation is from Henry K. Beecher, "Ethics and Clinical Research," *New England Journal of Medicine* 274 (1966): 1354–60. On the interpretation of Beecher as a proponent of a virtue account in the literature of medical ethics, see the brief statement in Mark Israel, *Research Ethics and Integrity for Social Scientists: Beyond Regulatory Compliance*, 2nd ed. (Los Angeles: Sage, 2015), p. 15.
69. 69. Thomas Keneally, *Schindler's List* (New York: Penguin Books, 1983), pp. 176–80.
70. 70. Rosalind Hursthouse, *On Virtue Ethics* (Oxford: Oxford University Press, 2001), p. 17. Other proponents of virtue ethics also stress that and how the virtues can guide action. See, for example, Julia Annas, "Why Virtue Ethics Does Not Have a Problem with Right Action," *Oxford Studies in Normative Ethics* 4 (2014): 13–33, and Annas, "Learning Virtue Rules: The Issue of Thick Concepts," in *Developing the Virtues: Integrating Perspectives*, ed. Annas, Darcia Narvaez, and Nancy E. Snow (New York: Oxford University Press, 2016), pp. 224–34. For an analysis of major positions in virtue ethics on this matter, see Liezl van Zyl, "Virtue Ethics and Right Action," in *The Cambridge Companion to Virtue Ethics*, ed. Russell, pp. 171–96.
71. 71. See Rosalind Hursthouse, "Virtue Ethics and the Treatment of Animals," in *Oxford Handbook of Animal Ethics*, ed. Beauchamp and Frey (2011), pp. 126–27; and Hursthouse, "Virtue Ethics," in *The*

Stanford Encyclopedia of Philosophy (Winter 2016 Edition), ed. Edward N. Zalta, available at https://plato.stanford.edu/entries/ethics-virtue/ (accessed April 11, 2018). See also Christine Swanton, *Virtue Ethics: A Pluralistic View* (New York: Oxford University Press, 2003), part 4; and Rebecca L. Walker and Philip J. Ivanhoe, eds., *Working Virtue: Virtue Ethics and Contemporary Moral Problems* (New York: Oxford University Press, 2009).

72. 72. See related reflections on the virtues, to which we are indebted, in Annas, *Intelligent Virtue*, chap. 3, esp. pp. 32–40.
73. 73. For a notably different analysis of "the link between virtues, principles, and duties" in biomedical ethics, see Edmund Pellegrino and David Thomasma, *The Virtues in Medical Practice* (New York: Oxford University Press, 1993), chap. 2. See also Pellegrino, "Professing Medicine, Virtue Based Ethics, and the Retrieval of Professionalism," in *Working Virtue*, ed. Walker and Ivanhoe, pp. 61–85. Other examinations of virtue ethics specifically in the context of health care include Rebecca L. Walker's "Virtue Ethics and Medicine," in *The Routledge Companion to Virtue Ethics*, ed. Besser-Jones and Slote, pp. 515–28; Justin Oakley, "Virtue Ethics and Bioethics," in *The Cambridge Companion to Virtue Ethics*, ed. Russell, pp. 197–220; Alan E. Armstrong, *Nursing Ethics: A Virtue-based Approach* (Houndmills, UK: Palgrave Macmillan, 2007); and literature discussed and cited in Chapter 2 of this volume.
74. 74. While widespread, the language of "consenting a patient" or "consenting a research subject" is objectionable because only the patient or subject can consent—health professionals and investigators do not "consent" patients or subjects; they provide the opportunity for and enable patients' and subjects' consent, for example, by providing relevant information. See our discussion of informed consent in Chapter 4, pp. 188–23.
75. 75. Brandt, "Toward a Credible Form of Utilitarianism," p. 166.

10

Method and Moral Justification

Can we justify moral conclusions in biomedical ethics? If we can, which methods can we legitimately and effectively use? The literature of the field offers numerous answers to these questions. In this chapter we step back from the first-order problems of normative and biomedical ethics that have largely preoccupied us to this point and reflect on second-order problems of method and justification. We assess the leading methods and forms of justification and defend an account that descends from John Rawls's justly celebrated theory of reflective equilibrium.

The first three sections of this chapter explicate and evaluate three models of method and justification. We then assess the arguments of critics of our methods and framework of principles. Finally, we connect our account of method and justification to the theory of common morality introduced in Chapter 1, the account of moral character developed in Chapter 2, and the analysis of moral status in Chapter 3.

Justification in Ethics

Justification has several meanings, some specific to disciplines. In law, justification is a demonstration in court that one has legally sufficient reasons and evidence for one's claim or for what one has been called to answer. In ethical discourse, the objective is to establish one's case by presenting morally sufficient reasons for it. A mere listing of reasons will not suffice because those reasons may not adequately support the conclusion that needs to be justified. Not all reasons are good reasons, and not all good reasons are sufficient for justification. We need to distinguish a reason's *relevance* to a moral judgment from its *sufficiency* to support that judgment and also to distinguish an *attempted* justification from a *successful* justification. For example, chemical companies in the United States at one time argued that the presence of toxic chemicals in a work environment provides a legally and morally sound reason to exclude women of childbearing age from a hazardous workplace, but the US Supreme Court overturned these policies on grounds that they discriminate against women.[1] The dangers to health and life presented by hazardous chemicals constitute a *good* reason for protecting employees from a workplace, but this reason is not a *sufficient* reason for a ban that impacts only women.

Several models of method and justification operate in normative ethical theory and contemporary biomedical ethics. We will analyze three models. The first approaches justification and method from a top-down perspective that emphasizes moral norms, as discussed in Chapter 1, and ethical theory, as discussed in Chapter 9. The second model approaches justification and method from a bottom-up perspective that emphasizes precedent cases, moral traditions, experience, and particular circumstances. The third does not assign priority to either a top-down or a bottom-up strategy. It emphasizes considered moral judgments and the overall coherence of facts, moral norms, and moral beliefs in a general moral framework. We defend a version of the third.

Top-Down Models: Theory and Application

A top-down model holds that we reach justified moral judgments through a structure of general normative precepts that support the judgments. This model is inspired by disciplines such as mathematics, in which a claim follows logically (deductively) from a credible set of premises. Justification occurs if and only if general principles or rules, together with the relevant facts of a situation, support an inference to a justified judgment. This model conforms to the way many people have been raised to think morally: It involves applying a general norm (principle, rule, ideal, right, etc.) to a clear case falling under the norm. The deductive form is sometimes considered an application of general precepts to particular cases, a conception that has encouraged use of the term *applied ethics*, a term we challenged but also used in Chapter 1.

The following is the deductive form involved in "applying" a norm (here using what is obligatory, rather than what is permitted or prohibited, although the deductive model can be fashioned for all three):

1. 1. Every act of description *A* is obligatory.
2. 2. Act *b* is of description *A*.

Therefore,

1. 3. Act *b* is obligatory.

A simple example is this:

1. **1x.** Every act in a patient's overall best interest is obligatory for the patient's doctor.
2. **2x.** Act of resuscitation *b* is in this patient's overall best interest.

Therefore,

1. **3x.** Act of resuscitation *b* is obligatory for this patient's doctor.

Covering precepts, such as 1 and 1x, occur at various levels of generality, but they are always universal in their logical form. The level of generality varies according to the specificity of the description A, while the statement's universal form is ensured by the claim that every act of such a description is obligatory. Particular judgments or beliefs are justified by bringing them under the scope of one or more moral rules; and the rules may be justified by bringing them under general principles, which in turn might be justified by appeal to a normative ethical theory. Consider a nurse who refuses to assist in an abortion procedure. The nurse might attempt to justify the act of refusal by the rule that it is wrong to kill a human being intentionally. If pressed, the nurse may justify this moral rule by reference to a principle of the sanctity of human life. Finally, the particular judgment, rule, and principle might all find support in an ethical theory of the sort discussed in Chapter 9 and in a theory of moral status of the sort discussed in Chapter 3.

This model functions smoothly in the straightforward case of a judgment brought directly and unambiguously under a rule or a principle—for example, "You must tell Mr. Sanford that he has cancer and will probably die soon, because a clinician must observe rules of truthfulness in order to properly respect the autonomy of patients." The top-down model proposes that the judgment, "You should not lie to Mr. Sanford," descends in moral content directly from the covering principle, "You should respect the autonomy of patients," from which we derive the covering rule, "You should not lie to patients."

Problems in the Model

This model suggests an ordering in which general theories, principles, rules, and rights enjoy moral priority over traditional practices, institutional rules, and case judgments. Although much in the moral life conforms roughly to this covering-norm conception, much does not. Particular moral judgments in difficult cases almost always require that we specify and balance norms (as discussed in Chapter 1), not merely that we bring a particular instance under a preexisting covering rule or principle. The abstract rules and principles in moral theories are extensively indeterminate. That is, the content of these rules and principles is too abstract in many situations to determine the specific acts that we should and should not perform. In the process of specifying and balancing norms and in making particular judgments, we often must take into account facts, cultural expectations, anticipated outcomes, and precedents to help assign relative weights to rules, principles, and theories.

The moral life often requires much more than general and specified norms. No general or specified norm (principle or rule) may clearly apply in a situation. The facts of cases are usually complex, and the different moral norms that can be brought to bear on the facts may yield inconclusive, or even contradictory, results. For example, in the controversy over whether it is permissible to destroy a human embryo in a petri dish for purposes of scientific research, embryo destruction does not clearly violate rules against killing or murder, nor does the rule that a person has a right to protect his or her bodily integrity and property clearly apply to the

destruction of extracorporeal human embryos. Even if all relevant facts are available, our selection of pertinent facts and pertinent rules may generate a judgment that is incompatible with another person's selection of facts and rules. Selecting the correct set of facts and bringing the right set of rules to bear on these facts is not reducible to a straightforward process of deduction.

The top-down model also creates a potentially infinite regress of justification—a never-ending demand for final justification—because each level of appeal to a covering precept requires a higher level to justify that precept. In theory, this problem could be handled by presenting a principle that is self-justifying or irrational not to hold, but proving that some principles occupy this status and that they justify all other principles or rules is an arduous demand that current ethical theory cannot meet. If all standards are unjustified until brought under a justified covering precept, it would appear, on the assumptions of this approach, that there are no justified principles or judgments.

A Theory of "Morality as a Public System"

One important version of top-down theory, and the most thoroughly examined in biomedical ethics (though it is not a pure deductivism), is the theory of Bernard Gert, as developed in bioethics with coauthors Danner Clouser and Charles Culver. Gert refers to his basic ethical theory as a theory of "morality as a public system." The public moral system is the institution of morality at work in our daily lives—that is, lived, pretheoretical morality—whereas a moral theory that describes and defends the norms of morality is a philosophical account. This theory can be thought of as top-down in the sense that the major elements of the theory are general moral rules, moral ideals, the morally relevant features of situations, and procedures for dealing with conflicts and assessing whether certain violations of moral rules are justified. Morality is envisaged in this theory as a public system of norms that are applicable to all persons in all places and times.[2]

When challenges arose to our framework of principles in the 1980s, the challengers who were grounded in Gert's moral theory emerged as our most unsparing critics in several articles and parts of books that expressed grave concerns about our prima facie principles. They coined the label "principlism" to refer to any account of ethics comprising a plurality of potentially conflicting prima facie principles. In our view, Gert and colleagues are not as distant from the views we defend in this book as they suppose. Like us, they understand the common morality as universal morality that is not relative to cultures, individuals, religions, or professional associations. However, Gert and colleagues reject both the language and the substance of our account of principles, while putting forward their account of impartial rules, moral ideals, and the definition of morality as a superior, alternative framework in biomedical ethics. We concentrate in the present section more on their criticisms of our principles and methods than on the nature and limits of top-down theories in general.[3]

First, Gert and colleagues charge that principles function as little more than names, checklists, or headings for values morally worth remembering while lacking deep moral substance and capacity to guide action. That is, principles do little more than point to moral themes that merit consideration by grouping those themes under broad headings of important moral notions. A second criticism is that because moral agents confronted with bioethical problems receive no specific, directive guidance from abstract principles, they are left free to deal with the problems in their own way, virtually as they wish. They may give a principle whatever interpretation and weight they wish, or even no weight at all. From this perspective, our account is insubstantial and permissive, largely because it lacks a controlling, comprehensive, tight theory. A third criticism is that the prima facie principles and other action guides in our framework often conflict, and our account is too indeterminate to provide a decision procedure to adjudicate the conflicts.

Clouser and Gert find these deficiencies especially obvious in the idea of principles of justice, as we examine accounts of justice in Chapter 7. They maintain that no specific guide to action derives from these principles and that all of the principles of justice we mention merely instruct persons to attend to matters of justice and think about justice, but they give no specific normative guidance about how to pursue the demands of justice. Because vagueness and generality underdetermine solutions to problems of justice, agents are free to decide what is just and unjust as they see fit.

Gert and Clouser also criticize our theory for giving status to beneficence as a principle of obligation. In a striking part of their theory, they maintain that there are no moral *obligations* of beneficence, although they hold that moral *ideals* of beneficence are vital parts of morality and should be encouraged. In their system, the only obligations in the moral life, apart from duties encountered in professional roles and specific stations of duty, are captured by moral rules that prohibit causing harm or evil—that is, rules of nonmaleficence. For Gert and colleagues, the general goal of morality is to minimize evil or harm, not to promote good. Rational persons can act impartially at all times in regard to all persons with the aim of not causing evil, but rational persons cannot impartially promote the good for all persons at all times.[4] This account ranks nonmaleficence as far more important than beneficence in the core dimensions of morality—a thesis we have rejected in previous chapters on grounds that the common morality accepts both of these as principles of obligation and does not prioritize either over the other.

The Limitations of "Morality as a Public System"

We agree that the problems Gert, Clouser, and Culver raise deserve sustained reflection, but we reject the key criticisms they direct at our account, some of which can be turned back on their theory. In particular, their criticism that our principles lack directive moral substance (as unspecified principles) applies to their rules in a near-identical way, because their rules are simply one level less abstract than our principles in the order of abstraction. Any norm, principle, or rule will have this problem if it is underspecified for the task at hand. All general norms, including Gert's moral rules, are designed to cover a broad range of circumstances. If general rules are not specified in biomedical ethics, they are almost always too general and will fail to provide adequate normative guidance. Like our principles, Clouser and Gert's rules (e.g., "Don't cheat," "Don't deceive," and "Do your duty") lack specificity in their original general form. One tier less abstract than our principles, their rules are in effect at the level of *partially specified* principles, which explains why their rules do, we agree, have a more directive and specific content than our more general principles. Our account of principles and rules does, however, include a set of moral rules similar to the rules embraced by Gert and his colleagues.[5] As discussed below, some of their rules presuppose or require our principles of respect for autonomy and beneficence, which they reject as basic moral norms of obligation.

Regarding their criticism that our principles are checklists or headings without deep moral substance, we agree that principles order, classify, and group moral norms that require additional content and specificity. However, until principles are analyzed and interpreted (as we do in every first section of Chapters 4–7) and then specified and connected to other norms (as we do in later sections of each of these chapters), it is unreasonable to expect more than a classification scheme that organizes the normative content and provides general rather than specific moral guidance.[6] Moreover, the balancing that is often required can occur only in concrete situations.

Regarding the Gert-Clouser criticism that principles conflict with other principles in ways that our account cannot handle, we acknowledge that moral frameworks of principles do not themselves resolve conflicts among principles and derivative rules. No framework of general guidelines could reasonably anticipate the full range of conflicts, but the Gert and Clouser system does no more to settle this problem than our framework does. In Chapter 1 we maintain that our theory handles this problem through both balancing and specification, whereas their account assumes that its "more concrete" rules escape the need for specification. Only a theory that could put enough content in its norms to escape conflicts and dilemmas in all contexts could live up to the Clouser-Gert demand. In our judgment, no general moral theory has ever achieved the goal of a fully specified system of norms for health care ethics.[7]

Experience and sound judgment are indispensable allies in resolving these problems. Thomas Nagel has forcefully argued that an unconnected heap of obligations and values is an ineradicable feature of morality, and W. D. Ross rightly argued that many philosophers have forced an architectonic of unwarranted simplicity on ethics.[8] Some critics of Ross's account and ours charge that we fail to achieve systematic unity in moral theory, but we regard some measure of disunity, conflict, and ambiguity as pervasive features of the moral life that are unlikely to be entirely eliminated by a moral theory. Moral theory offers suitable and effective methods such as

specification, balancing, and ways of adjusting norms to achieve consistency, but theories should not be expected to eliminate all untidiness, complexity, and conflict.

We accept the criticism that our principle-based analysis fails to provide a general ethical theory, but we do not consider it a telling objection. We do not claim to have constructed either a general ethical theory or a comprehensive theory of the common morality, and we do not claim that our principles and methods are analogous to or substitute for the principles and methods of justification in leading classical theories, such as utilitarianism, with its principle of utility, and Kantianism, with its categorical imperative. We express some skepticism about these general theories in Chapter 9 on grounds that the goal of a unified foundation for ethics is likely to misrepresent some vital aspects of the moral life.[9]

In response to the Gert-Clouser criticism that the principle of beneficence expresses a moral ideal, not a moral obligation, our claim is that their thesis distorts the common morality. Their theory holds that one is never morally required (except by role, professional, or community duties) to prevent or remove harm or evil, but only to avoid causing harm or evil. They recognize no requirement to *do* anything that confers a benefit or prevents a harm—only to *avoid* causing harms or harmful events and conditions.[10] Their thesis makes beneficence merely a moral ideal, and thereby misreads the commitments of the common morality, which requires some beneficent actions while recommending beneficent moral ideals.

The claim that beneficence is never morally required is not supported even within the heart of Gert's account of moral obligations, despite his statements to the contrary. In his book *Morality: Its Nature and Justification*, Gert relies in several places on the premise that one is morally obligated to act beneficently. For example, he interprets one of his ten basic moral rules, "Do your duty," to incorporate obligations of beneficence. Gert explains his system and its commitments as follows:

> Although duties, in general, go with offices, jobs, roles, etc., there are some duties that seem more general. … In any civilized society, if a child collapses in your arms, you have a duty to seek help. You cannot simply lay him out on the ground and walk away. In most civilized societies one has a duty to help when (1) one is in physical proximity to someone in need of help to avoid a serious evil, usually death or serious injury, (2) one is in a unique or close to unique position to provide that help, and (3) it would be relatively cost-free for one to provide that help.[11]

Gert maintains that all such requirements are supported "in any civilized society" by the foundational moral rule "Do your duty." These requirements are identical to the obligations that follow from beneficence, a term in wide use in ethical theory since at least the eighteenth century. Gert's duty to help is best understood as a specification of the general principle(s) of beneficence. In both his moral theory and ours this duty is not merely relative to the conditions of a "civilized society." It therefore is not the case that Gert's system lacks obligations of beneficence in our sense of the term.[12] To generalize, much in principlism that Clouser and Gert appear to reject is incorporated into or presupposed by their last unspecified rule, "Do your duty." Their theory of the moral system therefore does not provide an alternative to our substantive claims regarding the nature and scope of obligations.

Numerous substantive requirements of the common morality are also better expressed in the language of principles than in the language of rules. Consider the principle of respect for autonomy, which Gert and his colleagues find as problematic as principles of justice and beneficence. Their disregard of this principle renders their assessments of some cases convoluted and puzzling. Here is one such case: Following a serious accident, a patient, while still conscious, refuses a blood transfusion on religious grounds; he then falls unconscious, and his physicians believe that he will die unless he receives a transfusion. Gert and Culver argue that the provision of a blood transfusion under these circumstances is paternalistic and wrong because, after the patient regains consciousness following the transfusion, the physicians must then violate either the moral rule against deception or the moral rule against causing pain: If they did not tell the patient about the transfusion, they would violate the rule against deception; if they did tell him, they would cause him pain.[13]

Gert and Culver's rejection of the principle of respect for autonomy forces them through this convoluted process of reasoning to this problematic conclusion. Early on, their theory lacked the normative resources to argue that

the transfusion in this case is paternalistic and prima facie wrong because it violates the competent patient's expressed wishes and choices.[14] Initially, Gert's moral rule "Do not deprive of freedom" was narrowly construed to prohibit blocking a person's opportunities to take action. In order to address problems that arise from the blood transfusion case, and similar cases, Gert and his colleagues later interpreted this moral rule more broadly to also include the "freedom from being acted upon."[15] This expanded interpretation is reasonable, but their rule, so interpreted, then approximates the principle of respect for autonomy as we analyze it—a principle they say they reject.

We conclude that the top-down theory of moral rules developed by Gert and colleagues encounters several problems and lacks sufficient power to show that it is preferable to our principlist account. Nonetheless, it offers important insights into the moral life, to which we return later in this chapter.

BOTTOM-UP MODELS: CASES AND ANALOGICAL REASONING

Some writers in biomedical ethics concentrate on practical decision making without regard to general principles and theories. They believe that moral justification proceeds bottom up (inductively) by contrast to top-down (deductively). Inductivists, as we will refer to them, argue that we reason from particular instances to general statements or positions. For example, we use existing social practices, insight-producing novel cases, and comparative case analysis as the starting points from which to make decisions in particular cases and then to generalize to important moral norms. Inductivists emphasize an evolving moral life that reflects experience with difficult cases, analogy from prior practice, and exemplary lives and narratives of the sort we discussed in Chapter 2. "Inductivism" and "bottom-up models" are broad categories containing several methodologies that are wary of top-down theories. Pragmatism,[16] particularism,[17] and narrative approaches,[18] as well as some forms of feminism and virtue theory (as discussed in Chapter 2), arguably qualify as bottom-up accounts.

Inductivists propose that particular judgments in concrete cases provide warrants to accept moral conclusions independently of general norms. They see rules and principles as derivative, rather than primary, in the order of knowledge and of justification. Hence, the meaning, function, and weight of a principle derive from previous moral struggles and reflection. For example, physicians once regarded withdrawing lifesaving medical technologies from patients as an act of impermissible killing. After confronting agonizing cases, they and society came to frame many of these acts as cases of permissible allowing to die and sometimes as morally required acts of acknowledging treatment refusals by patients. This change resulted from extensive experience with cases of both withdrawing and declining to withdraw treatment. From this perspective, all specific moral norms arise and are refined over time; they never become more than provisionally secure points in a cultural matrix of guidelines.

Consider an example from the explosion of interest in surrogate decision making starting in the last quarter of the twentieth century. A series of cases, beginning with the influential case of Karen Ann Quinlan (1976),[19] challenged medical ethics and the courts to develop a new framework of substantive rules for responsible surrogate decision making about life-sustaining treatments, as well as authority rules regarding who should make those decisions. This approach involved a laborious examination of analogous cases, and testing new hypotheses against preexisting norms. Cases subsequent to *Quinlan* were addressed by appealing to similarities and dissimilarities to *Quinlan* and related cases. A string of cases with some similar features established the terms of the ethics of surrogate decision making over a course of several years.

Casuistry: Case-Based Reasoning

Proponents of casuistry, an influential version of bottom-up thinking in biomedical ethics, have revived a model that enjoyed an impressive influence in medieval and early modern philosophy and have refashioned it for modern biomedical ethics.[20] The term *casuistry* (from the Latin *casus*, meaning "case") refers to the use of case comparison and analogy to reach moral conclusions.[21]

Albert Jonsen and Stephen Toulmin, the two most prominent proponents of this approach, have expressed reservations about our framework of principles.[22] In general, casuists are skeptical of rules, rights, and general theories that are divorced from or developed independently of cases, history, precedents, and circumstances. Appropriate moral judgments occur, they argue, through an intimate acquaintance with particular situations and the historical record of moral judgments about similar cases. Casuists disavow the goal of a tidy, unified theory containing *inflexible* universal principles.[23]

However, casuists do not entirely exclude rules and principles from moral thinking, and they welcome them when they are consistent with their form of case analysis. Because we do not in this book accept inflexible, unbending principles, Jonsen and Toulmin are open to our interpretation and use of principles even though they think it does not penetrate to the core territory of moral thinking. As casuists, they insist that moral judgments are often made when no appeal to principles is available. For example, we make moral judgments when principles, rules, or rights conflict and no further recourse to a higher principle, rule, or right is available. Furthermore, when principles are interpreted inflexibly, irrespective of the nuances of the case, casuists see a "tyranny of principles"[24] in which attempts to resolve moral problems suffer from a gridlock of conflicting principles, and moral debate becomes both intemperate and interminable.

This impasse can be avoided, Jonsen and Toulmin argue, by focusing on points of shared agreement about cases rather than on shared principles. The following is their prime example, drawn from their experiences with the National Commission for the Protection of Human Subjects of Biomedical and Behavioral Research:

> The one thing [individual commissioners] could not agree on was *why* they agreed. … Instead of securely established universal principles, … giving them intellectual grounding for particular judgments about specific kinds of cases, it was the other way around.
>
> The *locus of certitude* in the commissioners' discussions … lay in a shared perception of what was specifically at stake in particular kinds of human situations. … That could never have been derived from the supposed theoretical certainty of the principles to which individual commissioners appealed in their personal accounts.[25]

In this account, casuistical reasoning rather than universal principles or rules forged agreement, even though commissioners accepted and explicitly stated that they reasoned from "established universal principles." According to Jonsen and Toulmin, the commissioners functioned successfully by appealing to paradigms and families of cases, despite the diverse principles and theoretical perspectives held by individual commissioners. Although commissioners cited moral principles to justify their collective conclusions—and unanimously endorsed several general principles in their *Belmont Report*[26]—Jonsen and Toulmin argue that these principles were in the end less important in the commission's moral deliberation than were judgments about cases.[27] They note specifically that the principlist approach endorsed in the *Belmont Report* came late in the work of the commission after it had deliberated about various types of cases, such as research involving prisoners and research on children.

A simple example illustrates the Jonsen-Toulmin claim that moral certitude resides in case judgments rather than principles or theory: We know that it is generally morally wrong to introduce significant risks to children in biomedical research that does not offer them the prospect of direct medical benefit. We are confident in the statement, "We should not give this healthy baby the flu in order to test a new decongestant," even though we may be unsure which principle controls this judgment or whether some viable theory sanctions it. The casuist's assessment is that we are almost always more secure in such particular moral conclusions than we are about the theory or principles that purport to show *why* these conclusions are correct. Practical knowledge about cases takes priority over theoretical knowledge. For example, if a principle or a theory instructed us to give the flu to children in order to test drugs, as some versions of utilitarianism seem to propose, this instruction would provide us with a good reason for rejecting that principle or theory. Moral certitude, then, is found at the bottom—that is, in judgments about particular cases, precedent cases, and practical reasoning—not at the top in theories, principles, or theoretical reasoning.

When confronting new cases, casuists compare them to paradigmatically right and wrong actions and to similar and acceptable cases, as well as to similar and unacceptable cases. Precedent cases and analogical reasoning are paramount in this method of moral thinking. If a new case arises involving a problem of medical confidentiality, casuists consider analogous cases in which breaches of confidentiality were justified or unjustified to see whether such a breach is justified in the new case. Paradigm cases become the enduring and authoritative sources of appeal. For example, the literature of biomedical ethics frequently invokes cases such as the Karen Ann Quinlan case and the Tuskegee syphilis experiments as sources of authority for new judgments. Decisions reached about moral rights and wrongs in pivotal cases become authoritative for new cases, and they profoundly affect prevailing standards of fairness, negligence, paternalistic interventions, and the like.[28]

A similar method appears in case law through the doctrine of precedent. When an appellate court decides a particular case, its judgment is positioned to become authoritative for other courts hearing cases with relevantly similar facts. Casuists argue that moral authority likewise develops from a social consensus about proper conduct that has been formed around cases. This consensus is then extended to new cases by analogy to the past cases around which the consensus was formed. As similar cases and similar conclusions evolve, a society becomes increasingly confident in its moral conclusions and acknowledges firm generalizations in the form of principles, rules, and rights in its evolving tradition of ethical reflection. These generalizations are interpreted as summary statements of a society's previously developed moral insights about cases.

The Limits of Casuistry

Casuists have sometimes overstated the power of their account of the nature of moral judgments and justification and have understated the value of competing accounts, but a balanced assessment of the role of cases in moral reasoning can remedy these problems. Some leading casuistical writers have already qualified their positions in ways that make them allies of, rather than competitors to, approaches that feature general principles.[29]

Casuists sometimes seem to suggest that paradigm cases speak for themselves or inform moral judgment by their facts alone, which is an implausible thesis. For the casuist to move constructively from case to case, a recognized and morally relevant norm must connect the cases. The norm is not part of the facts or narrative of the cases involved; it is a way of interpreting, evaluating, and linking cases. All analogical reasoning in casuistry requires a connecting norm to indicate that one sequence of events is morally like or unlike another sequence in relevant respects. The creation or discovery of these norms is not achieved merely by analogy. In addition to showing that one case is similar to another, casuists must show (1) that the two cases are similar in morally relevant respects, and (2) that the putatively paradigm cases, whether positive or negative, have moral authority.

Jonsen addresses this problem by distinguishing descriptive elements in a case from moral maxims embedded in the case: "These maxims provide the 'morals' of the story. For most cases of interest, there are several morals, because several maxims seem to conflict. The work of casuistry is to determine *which maxim* should *rule the case* and to what extent."[30] This thesis fits well with our views about prima facie principles and rules. Casuistry presupposes principles, rules, or maxims as essential moral elements in paradigm cases and in the assessment of new cases. As Jonsen concisely puts it, "The principles are, in the casuist's view, *embedded in* the [paradigm] case"[31] even though the principles are not reducible to the facts of the cases. Accordingly, the *paradigm case* must be understood as very different from the *facts* in that case. It is an amalgam of the facts and the morally relevant norms.

Basically, the casuists' paradigm cases combine both *facts* that can be generalized to other cases (e.g., "The patient refused the recommended treatment") and *settled values* that are generalized (e.g., "Competent patients have a right to refuse treatment"). These settled values are analytically distinct from the facts of particular cases. In casuistical appeals, values and facts are bound together in paradigm cases, and the central values are preserved from one case to the next. The more general the central values—the connecting norms—the closer they come in status to prima facie principles.

Casuists maintain that cases point beyond themselves and evolve into generalizations, but a key problem for casuists is that the cases can evolve in the wrong way if improperly handled from the outset. This problem of justification is worrisome because casuists have no methodological resource to prevent a biased development of case-based judgments or a neglect of morally relevant features of cases. A tyranny of the paradigm case can result, just as there can be a tyranny of unyielding principles or dogmas.

The casuists' approach to identifying and labeling cases often seems more intuitive than reasoned, with insufficient attention to the process of narrating cases. An obvious but fundamental point, emphasized by literary critics, is that a case is itself a mini-narrative. Presenting a case is telling a story, with context, circumstances, characters, conflicts, conduct, consequences, and so forth. Narration inevitably involves framing the story, selecting the details, and constructing the narrative. Accordingly, it is important to critically investigate the kind of evaluative and other assumptions that often, perhaps unwittingly, structure cases in particular ways and may lead to both classifications and conclusions that are not adequately examined or ultimately justified.[32]

Consider now the evaluative descriptions of two different cases. In a case reported in the *Journal of the American Medical Association* under the title "It's Over, Debbie,"[33] a medical resident injects a terminally ill woman with enough morphine to end her life in response to her request, uttered in their first encounter: "Let's get this over with." Jonsen classifies this case as one of killing—bringing it under a taxonomy of cases of killing, governed by various maxims—and then he reasons analogically from paradigm cases in this taxonomy.[34]

While Jonsen's case description, classification, and analysis are straightforward in this case, conflicts often arise about the evaluative judgments that are reached about cases because of the type and classification chosen. This problem is evident in another example, which we introduced in Chapter 5: the disconnection of a ventilator maintaining the life of a patient with amyotrophic lateral sclerosis (Lou Gehrig's disease). At a conference, physicians presented and described this as an end-of-life case, in which the "patient" decided to discontinue the ventilator.[35] However, the audience—many of whom were experienced in the long-term use of ventilators—challenged this description and classification. For them this was a "disability" case in which the patient needed better care, more complete information, and increased options, particularly to help him deal with the isolation he felt following his wife's recent death.

These disputes over narration and classification show the importance of examining the assumptions, perspectives, and evaluations that enter into the description of cases. The clinicians presenting this case thought it was a "textbook case" of decision making at the end of life, but the audience considered it "a story in which a life was ended as a result of failures of information and assistance by the presenters themselves."[36]

This example underlines the importance of attending to what John Arras categorizes as "moral diagnosis," that is, the process of determining what a case is centrally about. This analysis is required because "real life does not announce the nature of problems in advance."[37] In the absence of reliable pre-assigned categories and labels, it is essential to attend with imagination and discernment to the diagnostic process in cases in medical ethics just as in cases in medical care. This diagnostic process requires recognizing and reducing bias in the "describing, framing, selecting and comparing of cases and paradigms."[38] Bias-reduction strategies should include describing cases in fuller detail and narrating them from different perspectives, to be followed by careful analysis.

Because the casuistical method works only from the bottom up, it may lack the necessary critical distance from cultural blindness, rash analogy, and tyrannical popular opinion.[39] How is the casuist to identify unjust practices, predisposing bias, and prejudicial use of analogy to avoid one-sided judgments? Identification of the morally relevant features of a particular case depends on those who make judgments about cases, and these individuals may operate from unduly partial perspectives. In this regard, the ethics of casuistry contrasts sharply with a stable system of principles, human rights, and virtuous agents. Even if we are confident that morally mature cultures maintain resources for critical distancing and self-evaluation, these resources do not emerge from the methods of casuistry itself.

The root of the problem is that casuistry is a method that lacks and fails to provide content. As a vital instrument of thought, it displays the fundamental importance of case comparison and analogy in moral thinking, but it lacks initial moral premises, tools of criticism, and adequate forms of justification. It also lacks a substantive ground of "certitude," in Jonsen and Toulmin's language.[40]

Casuistry works best in showing that we appropriately reason by analogy and are often confident in the conclusions we reach. For example, if we feel better after using a certain medicine, we feel comfortable in recommending it to other persons, in the expectation that they too will feel better. A logical form is present in all uses of analogy: If some person or thing has one property associated with a second property, and another person or thing also has the first property, it is justified to infer that the second person or thing also has the second property. However, such analogies often fail: Our friends may not feel better after they take our favored medicine. Analogies never warrant a claim of truth, and we often do not know something by analogy that we think we know. The method of casuistry leaves us with this problem: No matter how many properties one case and a similar case share, our inference to another property in the second case may mislead or produce false statements.

These concerns do not amount to sufficient reasons for rejecting either the casuistical method or the use of analogy in moral reasoning. Both are helpful as long as we have a solid knowledge base that allows their proper use. However, to obtain that knowledge base, the casuistical method must be supplemented by norms of moral relevance that incorporate prior judgments of right and wrong conduct.[41] We will return to this problem of a proper knowledge base when we come to the subject of "considered judgments" later in this chapter.

Casuists sometimes confuse the fact that we may have no need for a general *ethical theory* for purposes of practical ethics with the lack of a need for *practical principles* and their specification. They also sometimes conflate certitude about principles with certitude about theory. One of our most important claims later in this chapter is that the general public and the mainstream of moral philosophy have found a "locus of certitude" in considered judgments about universal moral norms, without finding certitude in a particular moral theory about the foundation of these principles. We agree with casuists that in practical deliberation we often have a higher level of confidence in our judgments about particular cases than we have in appeals to moral theories, but the principles and rules that are central to the common morality enjoy the highest level of certitude.

In a major methodological statement, Jonsen describes connections between principles and casuistry:

> Principles, such as respect, beneficence, veracity, and so forth, are invoked necessarily and spontaneously in any serious moral discourse. … Moral terms and arguments are imbedded in every case, usually in the form of maxims or enthymemes. The more general principles are never far from these maxims and enthymemes and are often explicitly invoked. Thus, casuistry is not an alternative to principles, in the sense that one might be able to perform good casuistry without principles. In another sense, casuistry is an alternative to principles: they are alternative scholarly activities.[42]

While it is doubtful that *alternative* scholarly activities are at work, the two methods are different and complementary. The prima facie principles we propose are not vulnerable to the casuists' critique of rigid principles and are not excluded by their methodology. The movement from principles to specified rules is similar to Jonsen's account of casuistical method, which involves tailoring maxims to fit a case through progressive interactions with other relevant cases that are governed by maxims. Casuists and principlists should be able to agree that when they reflect on cases and policies, they rarely have in hand principles that were formulated without reference to experience with cases, or paradigm cases that lack embedded general principles.

AN INTEGRATED MODEL: REFLECTIVE EQUILIBRIUM

Accounts from "the top" (theories, principles, rules) and "the bottom" (cases, analogies, particular judgments) must be either supplemented or replaced in order to be practically useful in biomedical ethics. Neither general principles nor paradigm cases adequately guide the formation of justified moral judgments in many situations.

Instead of these models, we support a version of a third model of method, justification, and theory construction that is sometimes characterized as a coherence theory or coherentism. However, strict coherentists hold that there is no identifiable group of initial norms that are warranted beliefs, a thesis we reject. We hold that the framework of basic principles, their specifications, and balancing judgments provide critically important support beyond mere coherence.

Our account is heavily indebted to John Rawls's theory of *reflective equilibrium*, a term he coined to depict a way of bringing considered judgments, principles, and background theories into a state of equilibrium or harmony, that is, coherence.[43] Considered judgments are those least likely to be affected by conflicts of interest and other distorting influences and therefore appear at least provisionally acceptable on their own without argumentative support. We follow Rawls's lead by making the principles in our framework the main considered judgments at the roots of medical morality.

Method in ethics, in this account, begins with the considered judgments in which we have the highest confidence and believe to have the least bias. They are "judgments in which our moral capacities are most likely to be displayed without distortion." Examples are judgments about the wrongness of racial discrimination, religious intolerance, and political repression.[44] "Without distortion" does not refer to intuitively certain judgments. On what basis, then, can we be confident that our considered judgments are sufficiently free of bias and constitute acceptable starting points?

This problem is best handled by a delineation of the epistemic and moral qualities of persons or institutions that participate in the selection of considered judgments. Moral judges are entitled to claims to have reached considered judgments only if those judgments have been framed from a perspective that reins in conflicts of interest and other temptations of self-interest. And evaluators must exhibit absence of prejudice, relevant knowledge, and honesty, as well as attitudes of sympathy and compassion for the welfare of others. Evaluators must also display these attitudes in a consistent and sustained way. The point of appealing to these epistemic and moral virtues is to identify the conditions under which it is justified to claim that a judgment qualifies as "considered." Mere pervasiveness of some set of moral beliefs is not sufficient. Convergence reached by individuals *qualified to reach the considered judgments* is the essential condition.

Whenever some normative feature in a person's or a group's prevailing structure of moral views conflicts with one or more of their considered judgments (a contingent conflict), they must modify something in their viewpoint and strive to achieve equilibrium and overall coherence. Even the considered judgments that we accept as central in the web of our moral beliefs are subject to revision once a conflict occurs. The goal of reflective equilibrium is to match, prune, and adjust considered judgments, their specifications, and other relevant beliefs to render them coherent. We then must test the resultant guides to see if they yield incoherent results. If coherence proves impossible to achieve, we must readjust some points in the system of beliefs in a renewed search for coherence.

Many particular codes of ethics at work in medical institutions are coherent, but they also can be expected to produce problems of incoherence or incompleteness as the code is specified or modified. Our theory does not ask more than that agents faithfully specify and balance principles and then monitor the structure of norms with careful attention to overall coherence. Persons involved in rendering norms coherent should not expect an end to this process of revision that provides a complete normative account. Establishing policies and specifying norms in new directions using reflective equilibrium is a continuous work in progress—a relentless process of improving moral norms and increasing coherence.

Consider, as an example in medical morality, the place of the traditional moral axiom, "Put the patient's interests first." To hold together a system of beliefs in which this rule plays a critical role, we should seek to make the rule as coherent as possible with other considered judgments about responsibilities in clinical teaching, to subjects in the conduct of research, to patients' families, to sponsors in clinical trials, to health care insurance companies, to health care institutions such as hospitals, in public health, and so forth. The requirement to bring these diverse moral responsibilities into coherence and then test the results against other moral commitments is complex and daunting at times. It is difficult, if not impossible, to hold the intuitively attractive rule "Put the

patient's interests first" as absolute when we confront the many possible conflicts this rule can have with other commitments, including public and institutional policies. The rule is an acceptable starting premise—a considered judgment—but not an absolute principle. Accordingly, we are left with a range of options about how to specify this rule and then balance it against other norms to create an overall coherence of norms.[45]

A relatively simple example from the ethics of the distribution of organs for transplantation illustrates this moral problem. Policymakers have been attracted to each of the following options: (1) distribute organs by expected number of years of survival of transplant candidates in order to maximize the beneficial outcome of the procedure, and (2) distribute organs by using a waiting list to give every candidate an equal opportunity. As they stand, these two distributive rules are inconsistent. However, elements of both can be retained in devising a coherent policy by setting limits on these rules and specifying them into consistency. The outcome of this process must then be made coherent with all other relevant principles and rules, such as norms of nondiscrimination against the elderly and the role of patients' ability to pay in the allocation of expensive medical procedures.

The goal of bringing the overall set of considered judgments and related moral and factual beliefs into coherence is best understood as one version of "wide reflective equilibrium."[46] The goal of wide reflective equilibrium in our account is to identify the relevant particular judgments, rules, concepts, data, and theories as resources for moral reflection and to bring them into equilibrium or to modify or reject some of them as incoherent with the belief system. Moral views included are beliefs about particular cases, about rules and principles, about virtue and character, about consequentialist and nonconsequentialist forms of justification, about the role of moral sentiments, and the like.

Achieving a state of reflective equilibrium in which all beliefs fit together coherently, with no residual conflicts or incoherence, is an ideal that will be realized only partially. The trimming, repair, and reshaping of beliefs will need to occur repeatedly in response to new situations of conflicting norms. However, this ideal is not a utopian vision toward which no progress can be made. Particular moralities are works continuously under development, not finished products.[47]

As an example of the threat of incoherence in the search for reflective equilibrium, consider our limited support in Chapter 5 of physician-assisted death at a patient's request. We there take seriously slippery-slope arguments in opposition to physician-assisted dying, yet we support various forms of physician assistance. David DeGrazia has questioned our assertion that these two claims can be rendered consistent. He views our position as a "compromise [that] apparently leads to contradiction," and so to incoherence.[48] To see how the two views are consistent, not contradictory, we return to the distinction we introduced in Chapter 1 between the justification of policies and the justification of acts. Public laws sometimes justifiably prohibit conduct that might be morally justified in individual cases. In the instant case, two moral questions about physician-assisted hastening of death need to be distinguished: (1) Are physicians ever morally justified in complying with patients' requests for assistance in *acts* of hastened death? (2) Is there an adequate moral basis to justify the *legalization* of physician-assisted hastening of death? We argue in Chapter 5 that there are morally justified *acts* of assisting patients in hastening their deaths, but that once public considerations and consequences external to the private relationship between a physician and a patient are the issue—including the implications of legalized physician-assisted hastening of death for medical education and medical practice in hospitals and nursing homes—these external considerations may (but also may not) provide sufficient moral reasons for prohibiting physicians from engaging in such actions as a matter of public law. Accordingly, policies that legalize physician assistance are morally unacceptable under some circumstances, but acceptable under others. There is no inconsistency or incoherence in this position on physician-assisted hastening of death.

Justification is a matter of reflective equilibrium in our model, and it is never a matter of bare coherence because the body of substantive judgments and principles that cohere could be morally unsatisfactory. Bare coherence could be nothing more than a system of prejudices and therefore must be constrained by substantive norms. An example of this problem is the "Pirates' Creed of Ethics or Custom of the Brothers of the Coast."[49] Formed as a contract between marauders circa 1640, this creed is a coherent set of rules governing mutual assistance in

emergencies, penalties for prohibited acts, the distribution of spoils, modes of communication, compensation for injury, and "courts of honour" that resolve disputes. All crew members had to swear an oath of allegiance, commonly taken on a Bible. This body of substantive rules and principles, although coherent, is a moral outrage. Its requirement to bear arms for purpose of theft, acceptance of a distributive scheme of spoils, and provision of slaves as compensation for injury involve immoral practices. But what justifies us in saying that this code is a morally unacceptable code of ethics even if it is entirely coherent?

This question points to the importance of starting with considered judgments that are our most thoroughly examined moral beliefs. Once this collection is assembled, we need to cast the net more broadly in interpreting, specifying, and generalizing these beliefs. Certain normative views are unacceptable not merely because of incoherence. They are wrong because there is no way when starting from considered moral judgments that, through reflective equilibrium, we could wind up with anything approximating the provisions in the Pirates' Creed.

In this book we start in the process of reflective equilibrium with a set of considered judgments that are acceptable initially without argumentative support—in particular with the set of four principles as a framework for biomedical ethics. This kind of approach is commonly associated with foundationalist moral theories, whereas we have given a central role to coherence achieved through the process of reflective equilibrium. Coherence theory is widely considered anti-foundationalist, whereas our common-morality theory may seem to be inherently foundationalist. This is one reason we do not classify our account as a pure coherence theory, and instead present it as reaching for coherence after starting with considered judgments as the basic building blocks. Some philosophers will insist that "basic building blocks" entails a foundationalist account, but our theory is an attempt to retain the best from both foundationalist theory and coherentist theory.

We cannot here engage the many tangled issues about whether coherentism is philosophically preferable to foundationalism. Our path around these problems is to present an appropriately adjusted version of reflective equilibrium and to join it with our common-morality approach to considered judgments. In this way, coherence serves as a key constraint on the formulation, specification, and balancing of the norms that guide actions. This constraint cannot be compromised or avoided in the attempt to achieve justification when engaged in specification and balancing.

To avoid an unduly conservative or parochial set of beliefs, a wide body of moral experience should be consulted to discover points of convergence. Consider an analogy to eyewitnesses in a courtroom. If sufficient numbers of independent witnesses converge to agreement in recounting the facts of a story, the story gains credibility beyond the credibility of any one individual who tells it. This process helps eliminate biases or misperceptions found in some accounts and helps eradicate stories that do not converge and cannot be made consistent with the main lines of testimony. The greater the coherence in a story that descends from initially credible premises and convergent testimony, the more likely we are to believe it and accept the story as correct. As we increase the number of accounts, establish convergence, eliminate biased observations, and increase coherence, we become increasingly confident that our beliefs are justified and should be accepted. When we find wider and wider confirmation of hypotheses about what should be believed morally, the best explanation is that these hypotheses are the right ones, although other considerations will often be at work in determining whether moral beliefs held by the various parties truly have epistemic legitimacy.

In conclusion, we note some unresolved problems about the method of reflective equilibrium that we are not able to address here.[50] First, some ambiguity surrounds the precise aim of this method. It might be used in reflecting on public policies, constructing a moral philosophy, or strengthening an individual's set of moral beliefs. The focus might be on judgments, on policies, on cases, or on finding moral truth. Second, it can be difficult to determine when an effort to achieve reflective equilibrium is going well and to know when one has succeeded. Explicit uses of the method (by contrast to claims to be using it) are difficult to locate in the biomedical ethics literature.[51] Most discussions are theoretical and distant from contexts of practice. We are still learning how well or how poorly the method has served or can be improved to serve practical ethics. Third, the far-ranging objectives of truly wide reflective equilibrium are intimidating and may be unattainable ideals of

both comprehensiveness and coherence. The goal of bringing into coherence widely diverse sets of beliefs in wide reflective equilibrium has not yet been achieved.

COMMON-MORALITY THEORY

We return now to the section of Chapter 1 in which we first sketched our account of common morality as a source of considered judgments.[52] One of our assumptions is that no more central moral content exists as a starting point for biomedical ethics than the norms from which we have formulated our four clusters of principles. "No more central" should not be understood as an assertion that the principles provide the sole moral content. We do not claim that principles and derivative rules as we have formulated them establish the basic content of the common morality. We do not present the principles—or the principles together with the focal virtues and the human rights that we discuss—as alone constituting the common morality. The set of principles featured in Chapter 4 through Chapter 7 and our discussion of the moral virtues are *drawn from* the territory of common morality, however small or large it may be. Our thesis is merely that the principles are a reasonable formulation of some vital norms of the common morality and that the principles are well suited as a framework for biomedical ethics. We agree with some of our commentators that there is much more to the common morality than we capture in this book.[53]

All common-morality theories share several features: First, they rely on ordinary, shared moral beliefs for their starting content. Second, they hold that an ethical theory that cannot be made consistent with these pretheoretical moral values falls under suspicion. Third, all common-morality theories are pluralistic: They contain two or more nonabsolute (prima facie) moral norms.

Our common-morality theory does not view *customary* moralities as part of the common morality even though they may embody elements of the common morality. Our theory is committed to a global bioethics in the sense that the principles are universally applicable, not merely local, customary, or cultural rules. The general norms in the common morality provide a basis for evaluating and criticizing customary moral viewpoints when they are deficient. Our particular account of the common morality also unites it with the method of reflective equilibrium delineated earlier.

Some writers in ethical theory and applied ethics seem to think that we would justifiably have more confidence in our principles and considered judgments if we could justify them on the basis of a comprehensive ethical theory. However, this outlook has the cart pulling the donkey: We should have more confidence in an ethical theory if it could be shown coherent in a comprehensive way with the considered judgments including a variety of norms comprising the common morality. If an ethical theory were to reject the various central principles, rights, and virtues we have discussed, we would have a sound reason for skepticism about the theory rather than skepticism about these principles, rights, and virtues. Our presentation of principles, virtues, and rights, together with our attempts to show their consistency with other aspects of the moral life such as moral emotions, constitutes *the normative account* in this volume, but we do not claim that another common-morality theory could not be superior to our account.[54]

Moral Change

It is a matter of fact that particular moralities, customary practices, and so-called consensus moralities can and do change. They may even change by a complete reversal of position on some issues. For example, a code of research ethics might at one time endorse placebo-controlled trials only to condemn such trials at a later time. Our defense of the method of reflective equilibrium invites related questions about moral change. Considered judgments occupy a central position and are the root of many inferred beliefs, but even considered judgments are, in principle, revisable. It follows that no norm can, in the method we have defended, claim the privileged status of being immune to revision even though considered judgments are the central starting points.

However, both the method of reflective equilibrium and the fact of change in specified norms and particular moralities leave unresolved whether *the common morality*, which is a universal morality, can itself change by a

process of subtraction, addition, or substantive emendation. Moral change entails that what was not previously morally required (or prohibited) at a later time becomes morally required (or prohibited). Could it come to be the case, morally, that we no longer have to keep our promises, that we can lie and deceive, or that a vice can become a virtue? If such norms are not fixed, then they could evolve into different norms and alter the normative content of the domain of morality.

In Gert's theory, by contrast, change cannot occur in the norms of the common morality because the basic moral rules are essential and timeless: "A *general* moral rule concerns actions open to all rational persons in all societies at all times. … A general moral rule is unchanging and unchangeable; discovered rather than invented. … Since general moral rules apply to all rational persons at all times, obviously they cannot be invented, or changed, or subject to the will of anyone."[55] Gert's position is clear but overstated. To the extent we can envisage circumstances in which human society is better served by substantively changing or abandoning a norm in the common morality, change could occur and could conceivably be justified. For example, it is conceivable, however unlikely, that the rule that we must tell the truth could become so severely dangerous that we might therefore abandon the rule altogether. The possibility of such change, however unlikely, weakens the claim that there is a common morality with essential conditions and normative authority for all moral agents in all times and places.

It would be dogmatic to assert without argument that the basic norms of the common morality cannot change, but it is also difficult to construct a historical example of a central moral norm that has been or might be valid for a limited duration before being abandoned because a good moral reason supported its displacement. No evidence known to us suggests that societies have handled moral problems by either rejecting or altering basic norms in the common morality. As circumstances change, we find moral reasons for saying that a norm has new specifications or valid exceptions or can be outweighed by other norms. These adjustments are not reasons to discard the norm. To the contrary, they show the lengths we often go to retain basic norms.

Clear-cut exceptions exist to even the most indispensable rules, such as the rule against killing. Particular moralities have carefully constructed exceptions in cases of war, self-defense, criminal punishment, martyrdom, misadventure, and the like. There is no reason to think that we cannot continue to handle social change by allowing exceptions to one or more stable norms in the common morality. These exceptions can be made explicit through new specifications of principles or rules.

In at least one notable respect moral change in the application of norms in the common morality has occurred and will continue to occur. Even if abstract norms do not change, the *scope of their application* does change. That is, the set of individuals to whom many or all of these principles and rules are deemed to apply changes, and we may anticipate such changes to continue. Our arguments in Chapter 3 regarding moral status anticipate this problem: "Who qualifies as belonging to the moral community?" may be the same question as, "Who qualifies for moral status?" It is possible that we might radically alter our understanding of who or what qualifies for moral status, who should receive moral respect, and who qualifies for protection by the full range of basic rights. It is a matter of historical fact about the practice of slavery and the denial of rights to women that norms about scope have rightly changed in many societies.

We also can envision a situation in which corresponding rules would be *added* to the common morality, by contrast to rules being abandoned or swapped out. For example, the common morality could be expanded to include a rule of equal moral consideration of persons fashioned as a strict rule of nondiscrimination. Depending on the formulation of this rule, it could prohibit, for example, the various forms of discrimination that are now widely regarded in many customary moralities as tolerable and perhaps even thoroughly justified. Current examples are found in contexts that disallow women from serving as religious leaders; that allow discrimination against gay, lesbian, and transgender individuals; that allow small businesses to discriminate in hiring by choosing only persons of a preferred sex (e.g., an ethnic restaurant that hires only male servers); and so forth. Inclusion of a basic moral rule of equal consideration of persons that challenged such practices would constitute a substantial change in the common morality because no such rule now prevails.

This kind of change may currently seem unlikely to occur, but we can conceive of the conditions under which such changes would occur. Some might argue that the common morality has already been refined in a conspicuously similar manner by changes in the way enslaved people, women, people of differing ethnicities, people with disabilities, and persons from many groups who were once denied basic human rights have come to be acknowledged as owed equal moral consideration. These changes in the scope of the application of norms constitute major and actual, rather than hypothetical or merely conceivable, changes in moral beliefs and practices.

But are such historical changes, which have upgraded the moral status of various classes of individuals, truly changes in the common morality? The changes that have occurred in the way various classes of individuals such as racial groups, ethnic groups, and women are morally regarded seem to be changes in a recognition of moral status, changes in particular moralities, or changes in ethical and political theories, rather than changes in the common morality. It is plausible to hold that the common morality does not now, and has never, included a specific principle or provision of equal moral consideration for all individuals, whatever such a provision might entail. We are confident that empirical investigation of rules determining who should receive equal consideration would reveal pervasive differences across individuals and societies that are firmly committed to proper moral conduct. A theory of the common morality therefore should remain open to the possibility that the common morality could and should include rules of equal moral consideration for groups such as women, people of every race and ethnicity, people with disabilities, the great apes, and other parties now excluded. On the matter of equal moral consideration, the common morality is not where it should be, for reasons we provide in the following section.[56]

Finally, can we confidently assert that norms prohibiting practices such as owning slaves are justified by the common morality, even though these norms cannot be said to be themselves included in the common morality? It might be argued that the common morality has no explicit standards that prohibit slavery, even though such prohibitions are common in many societies as a specification of the common morality. From this perspective the common morality has the capacity to be mined by particular moralities to justify rules that prohibit slavery. The explicit commitments of the common morality to respect for autonomy, nonmaleficence, and the like contain implicit commitments to norms that, when coherently assembled, would prohibit practices such as slave-owning.

However, this position is conceptually unsatisfactory as an account of common morality because it leaves space for particular moralities not to specify in this way, which would leave these moralities open to the accommodation of practices of slavery. So understood, this position holds that slavery is not prohibited by the common morality. This interpretation of the principles in common morality fails to appreciate their conceptual and moral depth. Slave-owning clearly violates respect for autonomy and nonmaleficence, and introduction of a rule allowing this practice would leave the common morality in a state of moral incoherence, whether or not slave-owning societies recognize this fact. Given our theory that considered judgments provide the starting points for an account of common morality, an acceptance of respect for autonomy as a basic principle rules out owning another person, thoroughly controlling that person's actions, and the like. To think otherwise is to fail to understand the concept and principle of respect for autonomy. An antislavery provision is not merely a matter of specification of this principle. Rather, the prohibition of slavery is part of what it is to respect autonomy—and arguably to accept the principle of nonmaleficence as well. Slavery is not coherent with these principles and cannot be specified to be coherent with them. If slavery were allowed, there would always be an internal incoherence in the principles of the common morality. In short, enslaving human beings is ineradicably prohibited by basic commitments of common morality.

We will not further pursue this line of argument about the normative conceptual content of common morality, but we stress its importance. Changes in the scope of individuals or groups protected by the norms of the common morality are among the most momentous changes to occur in the history of moral practices. A *theory* of the common morality that denies our capacity to criticize and even condemn traditions, communities, groups, or individuals whose viewpoints are morally unacceptable would be an ineffectual and indefensible theory that misunderstands the moral depth of common morality. We discuss the philosophical basis of this claim in treating the second of three types of justification to be discussed in the following section.

Three Types of Justification of Claims about a Common Morality

Three methods of justification are available to justify claims about the common morality: (1) empirical justification, (2) normative theoretical justification, and (3) conceptual justification.[57] These three methods and their objectives have often been confused. Each type stands to justify a different conclusion or set of conclusions about the common morality. We do not present a justification that actually uses one or more of the three strategies—an immense project in the case of each of the three. Our sole aim in this section is to identify three available *types of justification* and what they would justify in a fully developed account.

Empirical justification. In Chapter 1 we stated that the existence of the common morality might be demonstrated empirically, although some skepticism exists about the prospects for achieving this goal. Some commentators have interpreted us as holding (in previous editions) that common-morality theory is empirical in nature and requires empirical proof.[58] This interpretation misses a range of diverse approaches we recommend to justify claims about the common morality; some of these approaches are empirical and some normative. We first consider what would be entailed in an empirical investigation into whether a common morality exists.

If an empirical investigation were to show that universally accepted norms are *in fact* discoverable in the moral beliefs of individuals, institutions, practices, and cultures, the claim that a common morality exists would be empirically justified by such a discovery. However, we have previously noted that multiple particular moralities exist and that similarities and differences between them are empirically confirmable. This claim about differences in particular moralities is noncontroversial. But we have hypothesized that some central norms of morality are held universally in common among morally committed persons. No available empirical studies throw into question whether some particular moralities accept, whereas other particular moralities reject, the norms of the common morality. Existing empirical data about moral beliefs generally descend from studies of particular moralities that were never designed to determine whether a universally accepted morality exists. These empirical investigations have usually studied cultural differences in the way moral rules have become embedded and applied in cultures and organizations, but they have not investigated whether a common morality exists. These studies succeed in showing cultural differences in the *interpretation*, *specification*, and *balancing* of moral norms, but they do not study or show that cultures accept, ignore, abandon, or reject the standards of the common morality. For example, empirical studies do not test whether a cultural morality rejects rules against theft, promise-breaking, or killing. Rather, investigators study what particular societies consider to be theft, promise-breaking, and killing; how they handle exceptive cases; and the like.

Some critics of our common-morality thesis claim that anthropological and historical evidence already speaks against the empirical hypothesis that is assumed in the claim that a universal morality exists.[59] However, these critics seem not to appreciate the nuances that surround the design of empirical research that would test hypotheses about the common morality. In principle, scientific research could either confirm or falsify the hypothesis that there exist a number of universal principles and rules of common morality. Such research would state which hypotheses are to be tested, how to formulate inclusion/exclusion criteria for study subjects, and why these hypotheses and criteria were selected. To date, critics who argue that available empirical studies falsify common-morality claims have not attended to the possibility that the common morality can be scientifically investigated, including by a testing of hypotheses.

The primary hypothesis that we propose for empirical testing is this: All persons committed to morality in their moral assessments accept at least the norms that we have claimed to be central to the common morality. The persons to be selected for inclusion in a study that investigates this hypothesis are (1) persons who pass a rigorous test of whether their beliefs include one identified considered moral judgment (to be stated in the study protocol), and (2) persons who are determined to take the moral point of view.[60]

We acknowledge that it would be difficult to design this empirical inquiry, but the goal is achievable despite problems of either missing the target (namely, the general moral beliefs of all and only those committed to morality) or begging the question by insisting on studying only persons known to accept what we have said about the common morality. The question could be begged either by (1) designing the study so that the only

persons tested are those who already have the commitments and beliefs the investigator is testing for (e.g., by presupposing our four clusters of principles) or (2) designing the study so that all persons are tested whether or not they are committed to moral norms. The first design risks biasing the study in favor of the hypothesis that a common morality exists. The second design risks biasing the study against this hypothesis.

These problems in research design are formidable, but not insurmountable. We have defined the *common morality* in terms of "the set of norms shared by all persons committed to morality." Some persons are committed to morality but do not always behave in accordance with their commitments; other persons are not committed to morality at all.[61] Persons who are not committed to morality do not fall within the scope of our claims and could not appropriately be included as subjects in an empirical study. Some might conclude that we have constructed a circular and self-justifying position. They might say that we are defining the common morality in terms of a certain moral commitment and then allowing only those who accept the norms we have identified to qualify as persons committed to morality. We appreciate that our position risks stipulating the content of "morality," but this risk should be manageable through careful research design. Here we provide only a basic outline of one design that would manage this risk and would allow the research to support or to falsify our hypothesis.

In a proper methodology, an investigation would include only persons who have already been screened to ensure that they are committed to *some one* norm of morality that is reasonable to expect all morally committed persons to accept. We suggest that a reasonable such principle is *the principle of nonmaleficence*, as it is unimaginable that any morally committed person would reject this general principle. Acceptance of this principle could serve as an inclusion criterion, and nonacceptance as an exclusion criterion. This choice of a single general norm does not bias the study because it does not preselect study subjects for their beliefs in any of the several other norms we have hypothesized to be central to the common morality. The group of persons to be tested would not be screened by presupposing any norm other than nonmaleficence. Persons not committed to the principle of nonmaleficence would be excluded from the group of study subjects.

The purpose of this study is to determine whether cultural or individual differences emerge in the included group over the acceptance of moral norms concerned with respect for autonomy, beneficence, justice, and other hypothesized norms in the common morality including promise-keeping, truth-telling, helping incompetent persons, respecting confidentiality, protecting severely vulnerable persons, and so forth. The research design could also test whether various norms are universally held that we have not even considered to be universal.

Should it turn out that the persons studied do not share the norms we hypothesize to be in the common morality, the research would show that there is no common morality of the sort we have envisioned, and our hypothesis would be falsified, or at least would require significant revision.[62] If norms other than the ones we have mentioned were demonstrated to be shared across the subjects tested, this finding would yield insights about the breadth of the common morality beyond our formulation.

Regarding the threat of circularity in our structuring of this research, we propose the following way to avoid it. If we limit the participants in the research to persons who have a moral commitment to the principle of nonmaleficence, then we could not assert that this empirical study could validly reach the conclusion that no moral norms whatever were found in common across cultures; the presupposed commitment to nonmaleficence ensures that this one norm is held in common among research participants. However, the methodology would not be fatally flawed by this finding. If no moral norms were found in this research to be held in common across cultures—other than the principle of nonmaleficence, which was presumed in the selection of subjects—then the general hypothesis that a common morality of the sort we have postulated would be thrown into serious doubt. To boldly generalize, what this study might validly show is either (1) that there is no common morality of the sort we have envisaged or (2) that some set of identifiable moral norms is held in common in addition to the assumed principle of nonmaleficence. If (2) were demonstrated, the identified set of norms would presumably constitute at least some part of the common morality (although perhaps not its complete set of norms).

We do not claim that empirical confirmation of the hypothesis that such a set of norms of common morality exists constitutes a *normative* justification of these norms of the common morality. Exclusively empirical

findings do not yield a normative justification. However, empirical findings can assist us in using the method of reflective equilibrium defended earlier in this chapter—in particular, "wide reflective equilibrium." There we were concerned with how to control for bias and lack of objectivity in the choice of considered judgments. One way to control for bias is to collect information about what is widely, preferably universally, agreed to be correct. This information could then be appropriately employed in attempts to reach reflective equilibrium. Disputed or unshared judgments are not well-positioned to qualify as considered judgments, whereas widely shared agreement is a relevant consideration. Shared agreement helps sustain claims about what qualifies as a body of considered judgments. Findings of universally shared agreement can, in this regard, be made integral to the justificatory process even though the collected information is empirical rather than normative.

Empirical information about commonly held norms can in this way contribute to the process of normative justification. With this caveat, we turn to nonempirical methods of normative theoretical justification.

Normative theoretical justification. Neither historical facts, such as facts about the history and traditions of medical ethics, nor social science facts of the sort envisaged in the previous section serve directly to justify moral norms. In Chapter 9 we discussed criteria of normative theories and the approach to justification taken by four different types of theory. Utilitarian theories, Kantian theories, rights theories, and virtue theories, among others, could be employed to provide a theoretical justification of the norms of the common morality. We argued that the norms supported in these theories tend to converge to the acceptance of the norms of the common morality, but, at the same time, that establishing this convergence does not amount to a moral justification. Establishing convergence in types of philosophical theories is an empirical demonstration, not a normative one.

What, then, can be said in support of a normative justification of common morality theory? Earlier in the present chapter we discussed Bernard Gert's attempts to justify the common morality in his books *Morality: Its Nature and Justification* and *Common Morality: Deciding What to Do*. Gert has shown that there is no reason why the norms in the common morality cannot be justified by a general ethical theory. We do not suggest that he has conclusively demonstrated that his particular ethical theory is correct; we claim only that he has shown that a normative ethical theory can be put to the work of justifying the norms of the common morality. Gert rightly says that his account of ethics does not make "*empirical* claims about morality," but rather provides justification of the substantive norms that constitute the common morality.[63]

In Gert's theory, common morality is justified on the basis of rationality. He regards it as clear to all rational persons that we should not act irrationally because irrational actions are those that should not be performed:

> Rational persons want to avoid death, pain, disability, loss of freedom, and loss of pleasure, and they know not only that they are fallible and vulnerable but that they can be deceived and harmed by other people. They know that if people do not act morally with regard to them, they will be at significantly increased risk of suffering some harm. If they use only rationally required beliefs, it would be irrational not to endorse common morality as the system to be adopted to govern the behavior of all moral agents.[64]

Acting irrationally bears a close relationship to acting in ways that will increase the likelihood of certain basic harms, and Gert argues that the goal of the moral rules is to prohibit causing these harms or contributing to conditions that cause them.[65]

Ethical theories other than Gert's and other than the four types discussed in Chapter 9 also might be employed to justify the common morality. For example, pragmatism is a type of theory that could be adapted to this purpose.[66] Pragmatic justification holds that moral norms are justified by their effectiveness in achieving the object of morality. Once we identify an operative purpose or objective of an institution or system of thought (in this case, the institution of morality), we can vindicate a set of standards if it is better for reaching the identified objectives than any alternative set of standards. For example, a pragmatist could hold that the goal or object of morality is to promote human flourishing by counteracting conditions that worsen the quality of people's lives and might argue that the norms of the common morality are the best instrument to combat these conditions. A set

of norms is pragmatically justified if and only if this set is the best way to promote human flourishing when all factors—including human limitations, shortcomings, and vulnerabilities—are taken into consideration.[67]

We will not here attempt a justification of the pertinent moral norms of the common morality by appeal to a particular type of general ethical theory such as Gert's or, alternatively, pragmatism, but we encourage these theoretical endeavors. Our aim in this brief subsection is modest. We have shown only that such theories have been and can be constructed and, if successful, they would normatively justify the norms of the common morality.

Conceptual justification.[68] In Chapter 1 we discussed the importance in metaethics of conceptual analyses of normative notions such as *right*, *obligation*, *virtue*, *justification*, and *responsibility*. The concept of *morality* is clearly connected to normative notions. We will here defend the view that the concept of morality contains normativity not only in the sense that this concept requires some action-guiding norms but also in the sense that it contains *certain specific* moral norms—that is, a body of norms in morality in the normative sense. No system of belief lacking these specific norms counts as morality, and if someone claimed that a system without these common-morality norms counts as morality, the claim should be rejected as conceptually mistaken.

Philippa Foot defends such a claim in a celebrated essay:

> A moral system seems necessarily to be one aimed at removing particular dangers and securing certain benefits, and it would follow that some things do and some do not count as objections to a line of conduct from a moral point of view. … [T]here are starting-points … fixed by the concept of morality. We might call them "definitional criteria" of moral good and evil, so long as it is clear that they belong to the concept of morality—to *the* definition and not to some definition which a man can choose for himself. What we say about such definitional criteria will be objectively true or false. …
>
> [I]t does not follow that we can settle all moral questions in this [definitional] way. … [T]he concept of morality while it fixes a great deal also leaves quite a lot open.[69]

We agree with Foot's position that certain norms are essential to the concept of morality and are starting-points, fixed by the concept, that state what is objectively correct. Critical principles, rules, rights, and virtues should be included as "fixing a great deal." These norms are essential to any system of *moral* norms. By contrast, some norms that are referred to as "moral," such as norms that reject human rights, are external to, and their content excluded by, the normative concept of morality, even though "morality" is commonly used in this way in both metaethics and the social and behavioral sciences. In the purely descriptive sense, "morality" refers to a group's codes of conduct or to individuals' important beliefs and attitudes about their conduct. There are plural descriptive "moralities," and their content and standards can differ greatly. However, accurate reports of morality in the descriptive sense have no implications for how all persons should behave, whereas in a normative sense of "morality" some actions are universally immoral and others universally morally required.[70]

The norms internal to morality in the normative sense are indispensable points of reference without which we could not get our moral bearings. As we have occasionally said about the four clusters of principles that provide the framework of norms in this book, they are starting points having a secure place in the common morality. One way of understanding this claim (and also a way of understanding our agreement with both Foot and Gert) is that these anchoring norms are crucial elements of the concept of morality, whereas the distinctive norms in particular moralities are not essential even though they might be entirely coherent with the common morality. By contrast, some "moralities" depicted in historical and social scientific literature may contain practices that contradict norms of morality in the normative sense—for example, a so-called medical morality of not reporting harmful medical errors to hospital administrators or patients. Such a "morality" is no more than a body of indefensible customs.

We do not claim that our four clusters of principles form the conceptual heart of the common morality in a way that other principles, rules, rights, and virtues do not. Our claim is merely that we *draw from* the common

morality to formulate the principles *of biomedical ethics* in our book. The two italicized parts of this sentence are critical: Unlike Gert, we do not claim to have removed the veil from the full set of norms that constitute the common morality. The norms in the common morality undoubtedly reach out beyond the principles and rules on which we concentrate. Put another way, we do not claim that our principles and rules exhaust the norms in the common morality. For instance, what we said in Chapters 1 and 2 about the virtues recognizes their secure place in the common morality.[71] Second, we claim for our framework of four clusters of principles only that these principles are well suited as a general framework of starting points for biomedical ethics. We do not claim more.

If this line of argument is cogent, it is a conceptual mistake, when using "morality" in the normative sense, to assert that morality allows persons to trade in slaves, coerce persons to be subjects in high-risk biomedical experimentation, or conceal harmful medical errors. The proposition that such practices are permissible might correctly characterize the beliefs of certain groups when "morality" is used in a descriptive sense, but it is conceptually incorrect of "morality" in the normative sense. Likewise, the proposition that "lying is always morally permissible" is an unacceptable general norm, even though "lying is not permissible" is only a prima facie rule that can sometimes be justifiably overridden. The fact that lying is sometimes justified does not entail that the rule "Lying is not permissible" is not a conceptually central norm of morality in the normative sense.

An example of these problems appears in the moral vices briefly mentioned in Chapters 1, 2, and 9, such as malevolence, dishonesty, lack of integrity, and cruelty. In morality in the normative sense these character traits are excluded from the domain of the morally acceptable, even though they too are not absolute vices. There may be rare circumstances in which dishonesty is appropriate, as with lying. Similarly, rules that disallow causing suffering to others are excluded. Each of these norms is prima facie rather than unconditionally wrong.

Adequate defense of these claims would require more extensive analysis of the concept of morality than we can undertake here. It would not be enough to argue that morality is the social institution that functions to ameliorate or counteract the tendency for things to go badly in human relationships. Morality would also have to be shown to be more than taking what some philosophers have called "the moral point of view"—that is, taking a view with a certain moral attitude such as compassion. These approaches have often not adequately been captured in the concept of morality in the normative sense.[72]

Moral pluralists may claim that there are multiple concepts of morality in the normative sense, but moral pluralism is a group-relative notion best interpreted as a version of "morality" in the descriptive sense. It would be incoherent to formulate the normative meaning of the term *morality* as consisting of the norms of multiple moralities, because contradictory advice would be given. One could deny that the term *morality* is univocal and then formulate two or more normative senses of *morality* (ns_1, ns_2, etc.), each with a different set of substantive norms, just as we can distinguish descriptive and normative senses. However, this maneuver is the functional equivalent of analyzing "morality" descriptively, not normatively.

In line with our arguments in Chapter 9 about convergence in moral theory, we caution against a vision of differences in moral theories as amounting to a pluralism of theory. These theoretical disagreements are usually about the foundations of morality. Theorists tend to assume the acceptability of, rather than disagree about, core moral norms such as not breaking promises, not harming others, and respecting autonomous choice.[73] Put another way, philosophers with different conceptions of the theoretical justification of universal morality tend not to disagree significantly on the substantive norms that comprise morality in the normative sense, despite their disagreements about theoretical foundations.

Problems for Common-Morality Theory

Our account of the common morality leaves unsettled problems that would have to be addressed to provide a complete account. Three questions deserve more attention than they can receive here.

Specification and judgment. First, do specified principles enable us to reach practical judgments, or are they too indeterminate to generate such judgments? Our theory requires that we specify to escape abstract

indeterminateness and reduce conflicts in order to provide more precise action-guiding content, but a danger exists of overspecifying a principle or rule, thereby leaving insufficient room for deliberation, judgment, and balancing of norms in some circumstances. Balancing judgments in concrete circumstances can be as important for moral thinking as specification.

However, without tighter controls on both permissible balancing and permissible specification, critics will say that too much room remains for moral judgments that are not adequately justified and yet are sanctioned or permitted by the theory. Questions that remain include, "Can the conditions intended to structure and constrain balancing that we presented in Chapter 1 reduce raw intuition to an acceptable level?" and "Can the constraints of our proposals about justification be tightened to respond to these concerns?"

Coherence in the common morality? We have linked reflective equilibrium to a common-morality theory and have attempted to integrate them as an approach to method and justification in ethics. But is it reasonable to expect that the common morality itself is coherent? If one argues, as we do, that a heap of obligations and values unconnected by a single first principle comprises the common morality, is it possible to show that morality is coherent (or is there a way of recasting it into coherence) without radically reconstructing norms so that they only vaguely resemble the norms that we claim in this book to be drawn from the common morality?

Theory construction. The language of "common-morality theory" suggests that an ethical *theory* can be constructed that is based only on norms drawn from the common morality. Is there good reason to believe that a theory—not merely a loosely connected collection of principles and rules—is possible? Perhaps general principles and rules, standards of moral virtues, and statements of human rights are all that we should aspire to, rather than a *theory* that conforms to the criteria of theories delineated at the beginning of Chapter 9. Perhaps "ethical theory" has been so diluted in meaning in the case of "common-morality theories" that we should abandon the goal of a theory.

In part, these problems turn on different expectations for a "theory." Gert and Clouser expect a strong measure of unity and systematic connection among rules and moral ideals, a clear pattern of justification, and a practical decision procedure that flows from a theory, whereas other philosophers are skeptical of one or more of these conditions, and even of the language of "theory."[74] We have encouraged moral theory in this chapter, as in Chapters 1, 2, and 9; but we have cautioned against expecting too much from ethical theories in the way of systematic tidiness and action guidance. No available ethical theory will eliminate the importance of specification, balancing, and reaching for reflective equilibrium as aids in practical ethics.

CONCLUSION

The model of working "down" by applying theories or principles to cases has attracted many who work in biomedical ethics, but we have argued that this model needs to be replaced or at least buttressed by the method of reflective equilibrium. We have also argued that we have reason to trust norms in the common morality more than the abstract norms often found in general theories. Ethical theories should not be expected to yield specific rules or judgments capable of resolving all contingent moral conflicts. No theory has such power. Nonetheless, we have not defended a so-called antitheoretical position. We have encouraged moral reflection of several types, including the development of moral theories as ways to discover and analyze the common morality and to determine the place of principles, rules, rights, and virtues in biomedical ethics.

Our theory of a framework of principles is committed to a global bioethics by presenting universally binding norms that constitute ineliminable starting points for determining what is ethically acceptable in all societies. This theory rejects the hypothesis that morality is ultimately reducible to local, customary, or cultural rules. It nonetheless remains sensitive to the need for particular moralities to be fashioned for groups with legitimate reasons for having their own specific rules for research and practice in medicine, health care, and public health.

NOTES

1. 1. U.S. Supreme Court, *United Automobile Workers v. Johnson Controls, Inc.*, 499 U.S. 187 (1991); argued October 10, 1990; decided March 20, 1991.
2. 2. K. Danner Clouser and Bernard Gert, "A Critique of Principlism," *Journal of Medicine and Philosophy* 15 (April 1990): 219–36. This article, and others that followed it, defend Gert's ethical theory as presented in his book *Morality: Its Nature and Justification*, 2nd rev. ed. (New York: Oxford University Press, 2005). See also Gert's *Common Morality: Deciding What to Do* (New York: Oxford University Press, 2004); and Gert, Charles M. Culver, and Clouser, *Bioethics: A Return to Fundamentals* (New York: Oxford University Press, 1997), and the second edition, retitled *Bioethics: A Systematic Approach* (New York: Oxford University Press, 2006). Both contain sustained criticisms of our views. However, Gert, Culver, and Clouser accept, in our judgment, both the language of the common morality and a conception of it less dissimilar to ours than might appear at first glance, even though the remaining differences are substantial. Their first publication specifically mentioning the subject of "common morality" was Clouser, "Common Morality as an Alternative to Principlism," *Kennedy Institute of Ethics Journal* 5 (1995): 219–36. See further Tom L. Beauchamp, "Principlism and Its Alleged Competitors," *Kennedy Institute of Ethics Journal* 5 (1995): 181–98; and Gert, Culver, and Clouser, "Common Morality versus Specified Principlism: Reply to Richardson," *Journal of Medicine and Philosophy* 25 (2000): 308–22. For critical assessments of Gert's moral theory, see Robert Audi and Walter Sinnott-Armstrong, eds., *Rationality, Rules, and Ideals: Critical Essays on Bernard Gert's Moral Theory* (Lanham, MD: Rowman & Littlefield, 2002); and Carson Strong, "Gert's Theory of Common Morality," *Metaphilosophy* 38 (2007): 535–45.
3. 3. By limiting our assessments we are unable to present the full dimensions of the top-down ethical theory developed by Gert, which he and his colleagues refer to as "morality as a public system." Gert's books present clear expositions of this theory. Gert has told us in private conversation that he takes "the whole public system" to constitute morality, and so to be within the scope of his moral theory. He emphasizes that he does not want his theory to be interpreted as reducible to normative statements of obligation (the moral rules) or top-down rules.

   For Gert and his colleagues' several discussions of principlism, see Clouser and Gert, "A Critique of Principlism"; Gert and Clouser, "Morality vs. Principlism," in *Principles of Health Care Ethics*, ed. Raanan Gillon and Ann Lloyd (Chichester, England: Wiley, 1994), pp. 251–66; Gert, Culver, and Clouser, *Bioethics: A Systematic Approach*, chap. 5; and Clouser and Gert, "Concerning Principlism and Its Defenders: Reply to Beauchamp and Veatch," in *Building Bioethics*: Conversations with Clouser and Friends on Medical Ethics, ed. Loretta M. Kopelman (Boston: Kluwer, 2002), pp. 183–99.

   For our view of the best way to understand the term "Principlism," see Tom L. Beauchamp and Oliver Rauprich, "Principlism," in *Encyclopedia of Global Bioethics*, ed. Henk ten Have (Switzerland: Springer Reference Series, 2016).

4. 4. Gert, Culver, and Clouser, *Bioethics: A Systematic Approach*, pp. 11–14, 32ff, passim.
5. 5. Gert and associates, like us, appeal to a relatively small number of norms drawn from the common morality. See Gert, Culver, and Clouser, *Bioethics: A Systematic Approach*, pp. 22–23, 34–36. We find their 9th and 10th rules to be too vague, general, and difficult to specify—roughly the problems they report they find about the principles in our theory.
6. 6. Gert has maintained in private conversation with us that once our principles are interpreted as normative headings under which rules fall, they become unobjectionable, but they also become expendable as mere classification systems. His published view is that "if specified principlism develops properly, it will become our account." See Gert, Culver, and Clouser, *Bioethics: A Return to Fundamentals*, p. 90. See also a clarification and partial retraction of their earlier criticisms of our position, in Clouser and Gert, "Concerning Principlism and Its Defenders: Reply to Beauchamp and Veatch," pp. 190–91.
7. 7. For a proposed method to handle this problem, see Gert, Culver, and Clouser, *Bioethics: A Systematic Approach*, pp. 27–32, 38–42, 83–87; and "Morality vs. Principlism," pp. 261–63. For relevant criticism of their claims, see Henry Richardson, "Specifying, Balancing, and Interpreting Bioethical Principles," *Journal of Medicine and Philosophy* 25 (2000): 285–307, esp. 293–97, which also appears in a revised version in *Belmont Revisited: Ethical Principles for Research with Human Subjects*, ed. James F.

Childress, Eric M. Meslin, and Harold T. Shapiro (Washington, DC: Georgetown University Press, 2005), pp. 205–27.

8. 8. Thomas Nagel, *Mortal Questions* (Cambridge: Cambridge University Press, 1979), pp. 128–37; and W. D. Ross, *The Right and the Good* (Oxford: Clarendon, 1930; reprinted Indianapolis, IN: Hackett, 1988). For a discussion of Ross's position, see David McNaughton, "An Unconnected Heap of Duties?" *The Philosophical Quarterly* 46, no. 185 (October 1996): 434–47.
9. 9. See, further, Michael Quante and Andreas Vieth, "Defending Principlism Well Understood," *Journal of Medicine and Philosophy* 27 (2002): 621–49.
10. 10. Gert, Culver, and Clouser, *Bioethics: A Systematic Approach*, pp. 11–13.
11. 11. Gert, *Morality: A New Justification of the Moral Rules* (New York: Oxford University Press, 1988), pp. 154–55.
12. 12. Cf. Gert, Culver, and Clouser, *Bioethics: A Systematic Approach*, pp. 89–93; and the formulation in Clouser and Gert, "Concerning Principlism and Its Defenders: Reply to Beauchamp and Veatch," pp. 190–91.
13. 13. See Gert and Culver, "The Justification of Paternalism," *Ethics* 89 (1979): 199–210; for additional critique, see James F. Childress, *Who Should Decide? Paternalism in Health Care* (New York: Oxford University Press, 1982), pp. 237–41. See also the later statements of their position on paternalism in Gert, Culver, and Clouser, *Bioethics: A Return to Fundamentals*, chap. 10, "Paternalism," and *Bioethics, A Systematic Approach*, chap. 10, "Paternalism and its Justification." These chapters differ in several respects.
14. 14. There are good reasons to believe that this paternalistic provision of a blood transfusion against a patient's oral directive, made while competent (if, indeed, he was competent), was *wrong*, but we set this complicated problem aside here. (See our discussion of paternalism in Chapter 6.)
15. 15. See, for example, Gert, Culver, and Clouser, *Bioethics: A Systematic Approach*, p. 36. A discussion of this shift in their interpretation of the rule not to deprive others of freedom, in part to address the Jehovah's Witness case and other cases, appears in their *Bioethics: A Return to Fundamentals*, p. 210, where they write: "In our previous discussion of the blood transfusion case … we did not realize that the doctor violated the patient's freedom, for we limited depriving of freedom to attempting to control behavior. We now realize that someone can deprive another of freedom by taking away control of what touches or goes into his body."
16. 16. See the comments, formulations, and frameworks in John D. Arras, "Pragmatism in Bioethics: Been There, Done That," *Social Philosophy and Policy* 19 (2002): 29–58; Arras, "Freestanding Pragmatism in Law and Bioethics," *Theoretical Medicine* 22 (2001): 69–85; these two essays by Arras have been incorporated with revisions into Arras, *Methods in Bioethics: The Way We Reason Now*, edited by James Childress and Matthew Adams (New York: Oxford University Press, 2017), chaps. 5 and 6; Henry Richardson, "Beyond Good and Right: Toward a Constructive Ethical Pragmatism," *Philosophy & Public Affairs* 24 (1995): 108–41; Joseph J. Fins, Franklin G. Miller, and Matthew D. Bacchetta, "Clinical Pragmatism: A Method of Moral Problem Solving," *Kennedy Institute of Ethics Journal* 7 (1997): 129–45; and Heike Schmidt-Felzmann, "Pragmatic Principles—Methodological Pragmatism in the Principle-Based Approach to Bioethics," *Journal of Medicine and Philosophy* 28 (2003): 581–96.
17. 17. See Alisa L. Carse, "Impartial Principle and Moral Context: Securing a Place for the Particular in Ethical Theory," *Journal of Medicine and Philosophy* 23 (1998): 153–69; Daniel Callahan, "Universalism & Particularism: Fighting to a Draw," *Hastings Center Report* 30 (2000): 37–44; and Earl Winkler, "Moral Philosophy and Bioethics: Contextualism vs. the Paradigm Theory," in *Philosophical Perspectives on Bioethics*, ed. L. W. Sumner and Joseph Boyle (Toronto: University of Toronto Press, 1996), pp. 50–78.
18. 18. For interpretations, defenses, and critiques of narrative approaches used in bioethics, see Rita Charon, "Narrative Medicine: A Model for Empathy, Reflection, Profession, and Trust," *JAMA: Journal of the American Medical Association* 286 (2001): 1897–1902; Charon, *Narrative Medicine: Honoring the Stories of Illness* (New York: Oxford University Press, 2006); Charon et al., *The Principles and Practice of Narrative Medicine* (New York: Oxford University Press, 2017), esp. chap. 5, "Deliver Us from Certainty: Training for Narrative Ethics" by Craig Irving and Rita Charon; Charon and Martha Montello, *Stories Matter: The Role of Narrative in Medical Ethics* (New York: Routledge, 2002); Hilde Lindemann Nelson, ed., *Stories and Their Limits: Narrative Approaches to Bioethics* (New York: Routledge, 1997), which includes Howard Brody, "Who Gets to Tell the Story? Narrative in Postmodern Bioethics," chap. 2,

and John Arras, "Nice Story, but So What? Narrative and Justification in Ethics," chap. 5, the latter of which appears in revised form in Arras, *Methods in Bioethics*, ed. Childress and Adams, chap. 4; Joan McCarthy, "Principlism or Narrative Ethics: Must We Choose between Them?," *Medical Humanities* 29 (2004): 65–71; and Anne Hudson Jones, "Narrative in Medical Ethics," *British Medical Journal* 318 (January 23, 1999): 253–56.

19. 19. *In the matter of Quinlan*, 70 N.J. 10, 355 A.2d 647, cert. denied, 429 U.S. 922 (1976).
20. 20. See Albert R. Jonsen and Stephen Toulmin, *The Abuse of Casuistry: A History of Moral Reasoning* (Berkeley: University of California Press, 1988); Baruch A. Brody, "A Historical Introduction to Jewish Casuistry on Suicide and Euthanasia," in Brody, ed., *Suicide and Euthanasia: Historical and Contemporary Themes* (Netherlands: Spring, 1989); John D. Arras, "Getting Down to Cases: The Revival of Casuistry in Bioethics," *Journal of Medicine and Philosophy* 16 (1991): 29–51, reprinted in revised form in Arras, *Methods in Bioethics*, ed. Childress and Adams, chap. 3; Carson Strong, "Specified Principlism: What Is It, and Does It Really Resolve Cases Better than Casuistry?" *Journal of Medicine and Philosophy* 25 (2000): 323–41; and Strong, "Critiques of Casuistry and Why They Are Mistaken," *Theoretical Medicine and Bioethics* 20 (1999): 395–411.
21. 21. Casuists have had relatively little to say about the nature or definition of a "case," or about the precise meaning of the term *casuistry*, but see Albert R. Jonsen, "Casuistry and Clinical Ethics," in *Methods in Medical Ethics*, 2nd ed., ed. Jeremy Sugarman and Daniel P. Sulmasy (Washington, DC: Georgetown University Press, 2010), pp. 110–11, 119; and Albert R. Jonsen, Mark Siegler, and William J. Winslade, *Clinical Ethics*, 8th ed. (New York: McGraw-Hill, 2015). The latter does not claim to be a work of casuistry, but its use of cases in clinical ethics is instructive.
22. 22. See, for example, the important statement by Jonsen in "Casuistry: An Alternative or Complement to Principles?" *Journal of the Kennedy Institute of Ethics* 5 (1995), esp. 246–47; see further Jonsen, "Strong on Specification," *Journal of Medicine and Philosophy* 25 (2000): 348–60, and Jonsen, "Morally Appreciated Circumstances: A Theoretical Problem for Casuistry," in *Philosophical Perspectives on Bioethics*, ed. Sumner and Boyle, pp. 37–49. See also the analysis and assessment of casuistry in James F. Childress, *Practical Reasoning in Bioethics* (Bloomington: Indiana University Press, 1997), chap. 2, "Ethical Theories, Principles, and Casuistry in Bioethics: An Interpretation and Defense of Principlism."
23. 23. The following are two major sources of claims in ethical theory to have a unified theory of the sort casuists presumably would disparage (these examples are ours, not ones selected by casuists): (1) Jeremy Bentham: "From utility then we may denominate a principle, that may serve to preside over and govern … several institutions or combinations of institutions that compose the matter of this science." *A Fragment on Government*, ed. J. H. Burns and H. L. A. Hart (London: Athlone Press, 1977), p. 416. (2) Henry Sidgwick: "Utilitarianism may be presented as [a] scientifically complete and systematically reflective form of th[e] regulation of conduct." *Methods of Ethics* (Indianapolis, IN: Hackett, 1981), bk. 4, chap. 3, § 1, p. 425.
24. 24. Stephen Toulmin, "The Tyranny of Principles," *Hastings Center Report* 11 (December 1981): 31–39. See further Toulmin's articles "How Medicine Saved the Life of Ethics," *Perspectives in Biology and Medicine* 25 (1982): 736–50; and "The Recovery of Practical Philosophy," *American Scholar* 57 (1988): 337–52. See additionally, at a high level of theory, F. M. (Frances Myrna) Kamm, *Bioethical Prescriptions: To Create, End, Choose, and Improve Lives* (Oxford: Oxford University Press, 2013).
25. 25. Jonsen and Toulmin, *Abuse of Casuistry*, pp. 16–19.
26. 26. See National Commission for the Protection of Human Subjects of Biomedical and Behavioral Research, *The Belmont Report: Ethical Principles and Guidelines for the Protection of Human Subjects of Research* (Washington, DC: DHEW Publication OS 78–0012, 1978); Childress, Meslin, and Shapiro, eds., *Belmont Revisited: Ethical Principles for Research with Human Subjects*; and Tom L. Beauchamp, *Standing on Principles: Collected Essays* (New York: Oxford University Press, 2010), chaps. 1–2.
27. 27. In addition to *Abuse of Casuistry*, see Jonsen's views on principles in "Casuistry and Clinical Ethics," pp. 112–18; and Toulmin, "The National Commission on Human Experimentation: Procedures and Outcomes," in *Scientific Controversies: Case Studies in the Resolution and Closure of Disputes in Science and Technology*, ed. H. Tristram Engelhardt, Jr., and Arthur Caplan (New York: Cambridge University Press, 1987), pp. 599–613.
28. 28. See Arras, "Getting Down to Cases: The Revival of Casuistry in Bioethics," pp. 31–33, reprinted in revised form in Arras, *Methods in Bioethics*, ed. Childress and Adams, pp. 45–51; and Jonsen and

Toulmin, *Abuse of Casuistry*, pp. 16–19, 66–67.

29. 29. Carson Strong is a striking example, beginning with his "Specified Principlism" in 2000, esp. p. 337. His theses now incorporate both principles and common morality; see especially his "Theoretical and Practical Problems with Wide Reflective Equilibrium in Bioethics," *Theoretical Medicine and Bioethics* 31 (2010): 123–40. Also notably accommodating to principles is Jonsen's "Casuistry and Clinical Ethics," esp. p. 120, where he says that casuistry "does not deny that certain methodological moves in moral theory might be quite relevant to casuistic thinking, such as the reflective equilibrium and specification methods." This accommodating, sometimes integrative, approach to casuistry and principles arguably began as early as Arras's 1991 article "Getting Down to Cases: The Revival of Casuistry in Bioethics," and was solidified in Jonsen's 1995 article, "Casuistry: An Alternative or Complement to Principles?" esp. pp. 248–49.
30. 30. Jonsen, "Casuistry as Methodology in Clinical Ethics," p. 298.
31. 31. Jonsen, "Casuistry and Clinical Ethics," p. 119.
32. 32. See, particularly, Tod Chambers, *The Fiction of Bioethics: Cases as Literary Texts* (New York: Routledge, 1999); and Chambers, "The Fiction of Bioethics: A Precís," *American Journal of Bioethics* 1, no. 1 (2001): 40–43, which focuses on and criticizes the ways many cases are used in bioethics. See, in response, James F. Childress, "Case Narratives and Moral Perspectives: An Appreciative Response to Chambers," *American Journal of Bioethics* 1, no. 1 (2001): 57–59, as well as responses by a number of other scholars in the same issue. See also Childress, "Narratives versus Norms: A Misplaced Debate in Bioethics?" in *Stories and Their Limits: Narrative Approaches to Bioethics*, ed. Nelson, chap. 17. For related problems of paradigm cases and ways to make casuistry effective in practice, see Annette Braunack-Meyer, "Casuistry as Bioethical Method: An Alternative Perspective," *Social Science and Medicine* 53 (2001): 71–81.
33. 33. Anonymous, "It's Over, Debbie," *Journal of the American Medical Association* 259, no. 2 (1988): 272.
34. 34. Jonsen, "Casuistry as Methodology in Clinical Ethics."
35. 35. J. K. Kaufert and T. Koch, "Disability or End-of-Life: Competing Narratives in Bioethics," *Theoretical Medicine and Bioethics* 24 (2003): 459–69.
36. 36. Kaufert and Koch, "Disability or End-of-Life," p. 462.
37. 37. Arras, "Getting Down to Cases."
38. 38. Loretta Kopelman, "Case Method and Casuistry: The Problem of Bias," *Theoretical Medicine* 15 (1994): 21–37, at 21.
39. 39. See Cass Sunstein, "On Analogical Reasoning," *Harvard Law Review* 106 (1993): 741–91, esp. 767–78; Kopelman, "Case Method and Casuistry"; Arras, "Getting Down to Cases"; Kevin Wildes, *Moral Acquaintances: Methodology in Bioethics* (Notre Dame, IN: University of Notre Dame, 2000), chaps. 3–4; and Mark G. Kuczewski, *Fragmentation and Consensus: Communitarian and Casuistic Bioethics* (Washington, DC: Georgetown University Press, 1997).
40. 40. For additional criticisms of casuistry, see Tom Tomlinson, *Methods in Medical Ethics: Critical Perspectives* (New York, Oxford University Press, 2012), chap. 4 ("Casuistry: Ruled by Cases"); and John Arras, "Theory and Bioethics," *Stanford Encyclopedia of Philosophy* (Winter 2016 Edition; first published 2010), ed. Edward N. Zalta, available at https://plato.stanford.edu/archives/win2016/entries/theory-bioethics/ (retrieved April 27, 2018).
41. 41. See John Arras, "A Case Approach," in *A Companion to Bioethics*, ed. Helga Kuhse and Peter Singer (Oxford: Blackwell, 1998), pp. 106–13, esp. 112–13.
42. 42. Jonsen, "Casuistry: An Alternative or Complement to Principles?" pp. 246–47.
43. 43. John Rawls, *A Theory of Justice* (Cambridge, MA: Harvard University Press, 1971; rev. ed., 1999), esp. pp. 20ff, 46–50, 579–80 (1999: 17ff, 40–45, 508–9). See also Rawls's comments on reflective equilibrium in *Political Liberalism* (New York: Columbia University Press, 1996), esp. pp. 8, 381, 384, and 399.
44. 44. Rawls, "The Independence of Moral Theory," *Proceedings and Addresses of the American Philosophical Association* 48 (1974–75): 8; and, more generally, Rawls, "Outline of a Decision Procedure for Ethics," *Philosophical Review* 60 (1951): 177–97.
45. 45. Compare the conclusions in Richardson, "Specifying, Balancing, and Interpreting Bioethical Principles," p. 302.

46. 46. Norman Daniels, "Wide Reflective Equilibrium in Practice," in *Philosophical Perspectives on Bioethics*, ed. Sumner and Boyle, pp. 96–114; Daniels, *Justice and Justification: Reflective Equilibrium in Theory and Practice* (New York: Cambridge University Press, 1996); Daniels, "Reflective Equilibrium," *Stanford Encyclopedia of Philosophy*, section 3, revision of October 2016 (first published April 28, 2003), available at https://plato.stanford.edu/entries/reflective-equilibrium/ (accessed March 21, 2018); Jeffrey Brand-Ballard, "Consistency, Common Morality, and Reflective Equilibrium," *Kennedy Institute of Ethics Journal* 13 (2003): 231–58; and Owen J. Flanagan, *The Geography of Morals: Varieties of Moral Possibility* (New York: Oxford University Press, 2017), pp. 123–27 (on "superwide reflective equilibrium").
47. 47. Rawls, *A Theory of Justice*, pp. 195–201 (rev. ed., 1999: 171–76).
48. 48. DeGrazia, "Common Morality, Coherence, and the Principles of Biomedical Ethics," *Kennedy Institute of Ethics Journal* 13 (2003): 219–30, esp. 226.
49. 49. Circa 1640. Published 1974 by Historical Documents Co., available at http://www.jollyrogercayman.com/web%20pages/pirates_creed.htm (accessed August 17, 2007).
50. 50. For these and related problems, see John D. Arras, "The Way We Reason Now: Reflective Equilibrium in Bioethics," in *The Oxford Handbook of Bioethics*, ed. Bonnie Steinbock (Oxford: Oxford University Press, 2007), pp. 46–71, reprinted with revisions as "One Method to Rule Them All? Reflective Equilibrium in Bioethics," in Arras, *Methods in Bioethics*, ed. Childress and Adams, chap. 8; Daniels, "Reflective Equilibrium" (2016 revision), section 4; Strong, "Theoretical and Practical Problems with Wide Reflective Equilibrium in Bioethics"; Michael R. DePaul, *Balance and Refinement: Beyond Coherence Models of Moral Inquiry* (London: Routledge, 1993); Kai Nielsen, "Relativism and Wide Reflective Equilibrium," *Monist* 76 (1993): 316–32; and David DeGrazia and Tom L. Beauchamp, "Philosophical Methods," in *Methods of Bioethics*, 2nd ed., ed. Sugarman and Sulmasy, pp. 37–53.
51. 51. For a work in biomedical ethics that makes a reasonable claim to be using the method of reflective equilibrium throughout the book, see Allen Buchanan, Dan W. Brock, Norman Daniels, and Daniel Wikler, *From Chance to Choice: Genetics and Justice* (Cambridge: Cambridge University Press, 2000).
52. 52. Revisions of our theory across the eight editions of this book have benefited along the way from the criticisms and constructive suggestions of Ruth Faden, Oliver Rauprich, John Arras, Allen Buchanan, Norman Daniels, Bernard Gert, Dan Clouser, Rebecca Kukla, Carson Strong, Albert Jonsen, Earl Winkler, Frank Chessa, Robert Veatch, David DeGrazia, Ronald Lindsay, Avi Craimer, Henry Richardson, Marta Dias Marcelos, Bettina Schöne-Seifert, and Michael Quante.
53. 53. See especially Rebecca Kukla, "Living with Pirates: Common Morality and Embodied Practice," *Cambridge Quarterly of Healthcare Ethics* 23 (2014): 75–85; and Oliver Rauprich, "Common Morality: Comment on Beauchamp and Childress," *Theoretical Medicine and Bioethics* 29 (2008): 43–71.
54. 54. We were criticized for an incautious formulation of this point by Jan Reinert Karlsen and Jan Helge Solbakk, "A Waste of Time: The Problem of Common Morality in Principles of Biomedical Ethics," *Journal of Medical Ethics* 37 (2011): 588–91.
55. 55. Gert, *Morality: Its Nature and Justification*, pp. 114–15; and Gert, Culver, and Clouser, *Bioethics: A Systematic Approach*, p. 104. See also the gloss in Gert, Culver, and Clouser, "Common Morality versus Specified Principlism: Reply to Richardson," pp. 310, 316.
56. 56. We are indebted in this formulation to Ronald A. Lindsay, "Slaves, Embryos, and Nonhuman Animals: Moral Status and the Limitations of Common Morality Theory," *Kennedy Institute of Ethics Journal* 15 (December 2005): 323–46.
57. 57. The ideas in this section draw on Beauchamp, *Standing on Principles: Collected Essays*, chapter 11.
58. 58. For sources that make such claims and the unlikely character of the claims, see Peter Herissone-Kelly, "The Principlist Approach to Bioethics, and Its Stormy Journey Overseas," in *Scratching the Surface of Bioethics*, ed. Matti Häyry and Tuija Takala (Amsterdam: Rodopi, 2003), pp. 65–77, esp. 66; Herissone-Kelly, "Determining the Common Morality's Norms in the sixth edition of *Principles of Biomedica*l Ethics," *Journal of Medical Ethics* 37 (2011): 584–87; Ronald A. Lindsay, "Bioethics Policies and the Compass of Common Morality," *Theoretical Medicine and Bioethics* 30 (2009): 31–43, first section; Rebecca Kukla, "Living with Pirates"; and William T. Branch, "Is Rorty's Neopragmatism the 'Real' Foundation of Medical Ethics: A Search for Foundational Principles," *Transactions of the American Clinical and Climatological Association* 117 (2006): 257–71. We profited from these criticisms and have attempted to remove some of the unclarities in our previous editions.

59. 59. See Leigh Turner, "Zones of Consensus and Zones of Conflict: Questioning the 'Common Morality' Presumption in Bioethics," *Kennedy Institute of Ethics Journal* 13 (2003), 193–218; Donald C. Ainslie, "Bioethics and the Problem of Pluralism," *Social Philosophy and Policy* 19 (2002): 1–28; Carson Strong, "Exploring Questions about Common Morality," *Theoretical Medicine and Bioethics* 30 (2009): 1–9; and DeGrazia, "Common Morality, Coherence, and the Principles of Biomedical Ethics."
60. 60. "The moral point of view" is a designation that descends from a number of moral theories first developed in the 1950s. The theory presents an ideal of moral judgment; the core of the idea is that the moral point of view is the one that would be taken by impartial, dispassionate, and disinterested judges. The most detailed work on the subject is Kurt Baier, *The Moral Point of View* (Ithaca, NY: Cornell University Press, 1958). On the history, scope, and influence of the theory, see Kai Nielsen, "Moral Point of View Theories," *Crítica: Revista Hispanoamericana de Filosofía* 31 (1999): 105–16.
61. 61. In saying that some persons are not committed to morality, we do not mean that they are not dedicated to a way of life that they consider a moral way of life or that anthropologists would say they are not committed to morality. Extreme religious fanatics and political zealots have this self-conception even as they act against or neglect the demands of the common morality.
62. 62. If the selected group shares the norms, this fact supports the idea of a common morality, but it is not conclusive. For a conclusive confirmation one would need to investigate all persons committed to a moral way of life, which is not feasible. Thus, there remains an issue of what constitutes sufficient evidence.
63. 63. Bernard Gert (and subsequently revised by Joshua Gert), "The Definition of Morality," *The Stanford Encyclopedia of Philosophy* (Fall 2017 Edition), ed. Edward N. Zalta, available at https://plato.stanford.edu/archives/fall2017/entries/morality-definition/ (accessed April 20, 2018) (italics added).
64. 64. Gert, *Common Morality: Deciding What to Do*, p. 84.
65. 65. Gert, *Morality: Its Nature and Justification*, pp. 29–33, 39–41, 181.
66. 66. For a pragmatic account with direct relevance to biomedical ethics and our claim about pragmatic justification, see Henry S. Richardson, *Articulating the Moral Community: Toward a Constructive Ethical Pragmatism* (New York: Oxford University Press, 2018).
67. 67. See further Tom L. Beauchamp, "A Defense of the Common Morality," *Kennedy Institute of Ethics Journal* 13 (2003): 259–74; Oliver Rauprich, "Common Morality: Comment on Beauchamp and Childress," pp. 43–71, at 68; Rauprich, "Specification and Other Methods for Determining Morally Relevant Facts," *Journal of Medical Ethics* 37 (2011): 592–96; and K. A. Wallace, "Common Morality and Moral Reform," *Theoretical Medicine and Bioethics* 30 (2009): 55–68.
68. 68. Revision of the account in this section has benefited from published criticisms by and private conversations with Peter Herissone-Kelly, Bernard Gert, Oliver Rauprich, and Rebecca Kukla. Herissone-Kelly appropriately, and constructively, criticized us in his original work on the subject, "The Principlist Approach to Bioethics."
69. 69. Foot, *Moral Dilemmas* (Oxford: Oxford University Press, 2002), pp. 6–7. Peter Herissone-Kelly directed us to this passage; see his use of it in "Determining the Common Morality's Norms in the Sixth Edition of *Principles of Biomedical Ethics*," *Journal of Medical Ethics* 37 (2011): 584–87, at 584.
70. 70. For a superb explication and defense of this distinction between the descriptive and the normative, see Gert's "The Definition of Morality."
71. 71. See Chapter 1, pp. 3–4, on central virtues ("ten examples of moral character traits, or virtues" recognized in the common morality).
72. 72. Philosophers who attempt to analyze the concept of morality exhaustively in terms of descriptive conditions miss what is morally most important in the concept. These theories often describe morality as composed of (1) norms that are regarded as supremely authoritative and of overriding social importance, or (2) norms that are prescriptive in form (i.e., action-guiding imperatives that do not describe states of affairs), or (3) norms that are universalizable, or (4) norms that harmonize pro and con interests, or (5) norms that require other regarding conduct, or (6) a combination of some of these five. By design, these accounts do not address whether there is a specific normative content that is privileged and constitutive of morality. Proponents of this form of theory, which is often concerned to distinguish moral judgments and norms from nonmoral ones, include John Hartland-Swann, *An Analysis of Morals* (London: George Allen & Unwin, 1960); William K. Frankena, "What Is Morality?" in his *Thinking about Morality* (Ann Arbor: University of Michigan Press, 1980), chap. 1; and Gerald Wallace and A. D. M. Walker, *The Definition of*

Morality (London: Methuen, 1970). See also the discussion in James F. Childress, "The Identification of Ethical Principles," *Journal of Religious Ethics* 5, no. 1 (1977): 39–68; the original version of this essay appears in the *Belmont Report: Ethical Principles and Guidelines for the Protection of Human Subjects of Research*, by the National Commission for the Protection of Human Subjects of Biomedical and Behavioral Research, DHEW Publication No. (OS) 78-0013, 1977, Appendix, vol. I.

73. 73. Gert defends a somewhat similar view in "The Definition of Morality."
74. 74. A relevant example of theory skepticism is Annette Baier, *Postures of the Mind* (Minneapolis: University of Minnesota Press, 1985), pp. 139–41, 206–17, 223–26, 232–37.

INDEX