Virtual Mentor
American Medical Association Journal of Ethics
July 2012, Volume 14, Number 7: 563-566.

HEALTH LAW

Informed Consent: What Must a Physician Disclose to a Patient?

Bryan Murray

Informed consent is at the heart of shared decision making—a recommended approach to medical treatment decision in which patients actively participate with their doctors. Patients must have adequate information if they are to play a significant role in making decisions that reflect their own values and preferences, and physicians play a key role as educators in this process.

Many patients may have a limited understanding of medicine, so it is difficult, if not impossible, for a physician to confirm that a patient has given adequately informed consent. Hence, it is almost self-evident that adherence to the doctrine of informed consent requires a physician to disclose enough about the risks and benefits of proposed treatments that the patient becomes sufficiently informed to participate in shared decision making. A practicing physician may find it difficult to strike a balance between too much and too little information. This article will discuss legal standards that define what types of risk and other information a physician must disclose in facilitating informed consent, as well as disclosures that are not legally required.

Informed Consent

The legal doctrine of informed consent can be traced back to the post-World War II Nuremburg Code, a set of guidelines drafted to ensure that unethical "medical" experiments were no longer carried out in the name of science. The doctrine is founded on the general principle that a person of the age of majority and sound mind has a legal right to determine what may be done to his or her body [1]. Thus, when a patient is subjected to a procedure he or she has not agreed to, the physician performing that procedure is violating the patient's legal rights and may be subject to medical malpractice litigation, removal from preferred-provider lists, or the loss of hospital privileges.

To avoid legal action, according to the doctrine of informed consent, physicians must disclose enough information for the patient to make an "informed" decision. However, because informed consent laws and principles do not specify the amount of information that must be disclosed, physicians might find it useful to know what they must typically disclose.

Traditionally, courts held that a physician's duty to disclose information to the patient depended upon community disclosure standards—whether the majority of physicians within a particular community would customarily make such a disclosure

[2]. More recently they have acknowledged problems with the community disclosure standard, chiefly that it creates an incentive for physicians to protect themselves by collectively limiting the standard disclosures, which is not in patients' best interests. In effort to address this problem, the D.C. Circuit Court of Appeals dramatically altered the physician's duty to disclose in the seminal case *Canterbury v. Spence* [3].

In *Canterbury*, a young man was advised by his physician to undergo a laminectomy in an effort to alleviate back pain. The physician, aware that 1 percent of laminectomies resulted in paralysis, did not advise the patient of the risk because he believed this might cause the patient to reject the useful treatment. Following the procedure, the patient fell from his hospital bed and was paralyzed. It remained uncertain whether the laminectomy procedure or the patient's fall caused the paralysis.

The patient sued, alleging that the physician failed to inform him of the risks associated with the procedure. The court held that "the standard measuring [physician] performance…is conduct which is reasonable under the circumstances" [3]. In other words, the court held that, instead of adhering to the community disclosure standard, *physicians are now required to disclose information if it is reasonable to do so*. Essentially, to establish true informed consent, a physician is now required to disclose all risks that might affect a patient's treatment decisions.

In *Canterbury*, the decision outlined key pieces of information that a physician must disclose:

> (1) condition being treated; (2) nature and character of the proposed treatment or surgical procedure; (3) anticipated results; (4) recognized possible alternative forms of treatment; and (5) recognized serious possible risks, complications, and anticipated benefits involved in the treatment or surgical procedure, as well as the recognized possible alternative forms of treatment, including non-treatment [4, 5].

In two informed consent cases following *Canterbury*, physicians have also been required to disclose (1) personal or economic interests that may influence their judgment (*Gates v. Jenson*) [6] and (2) all diagnostic tests that may rule out a possible condition (*Jandre v. Physicians Insurance Co of Wisconsin*) [7]. In *Arato v. Avedon*, however, physicians were not required to disclose particular statistical life expectancy rates to a patient suffering from pancreatic cancer, mainly on the grounds that statistics do not usefully relate to an individual's future [8].

The decision in *Nixdorf v. Hicken* stipulated that physicians must also disclose information that a reasonable person in the patient's position would find important [9]. In this case, a doctor left a surgical needle in his patient and was held to have a duty to disclose any information pertinent to the patient's treatment, including the patient's physical condition following treatment [9].

Similarly, a physician must also explain any benefits or risks that may be significant to the particular patient. For example, any risk of injury to a patient's hand is especially important to a concert violinist or professional baseball pitcher. In the briefest terms, a physician is required to provide general information about a proposed diagnosis or treatment and more personalized information about how the treatment might reasonably affect the particular patient.

Truly informed consent may also require disclosure of potential risks associated with not seeking treatment. In the California case *Truman v. Thomas*, in which a woman had refused a pap smear, the court held that a physician had a duty to disclose to her the possibility that precancerous cells might develop, uncaught, into cervical cancer if she declined to undergo the procedure [10].

Exceptions

While a physician is required to disclose all reasonable information, he or she is not required to disclose a risk that is not inherent in proper performance of the procedure—a risk, in other words, that would result only from the procedure's being performed incorrectly [11, 12].

The courts have noted two additional exceptions to the requirement that physicians elicit and document informed consent. The first applies when both (1) the patient is unconscious or otherwise incapable of consenting and (2) the benefit of treating the patient outweighs any potential harm of the treatment. Under these circumstances, the physician is not required to obtain informed consent before treating, but must do so as soon as it is medically possible [13, 14].

The second exception applies when disclosing medical information would pose a threat to the patient. If, for example, a patient has become so emotionally distraught that he or she would become incapable of making a rational decision, courts generally do not require disclosure [15]. If disclosure is likely to cause psychological harm to the patient, a physician does not have a duty to disclose [16]. However, a physician cannot use the exception to withhold information merely because he or she thinks the information may cause the patient to refuse a specific treatment. In other words, a physician must disclose information that a reasonable person would want to have for decision making, even though that information may cause the patient to refuse treatment that the physician believes is in the patient's best interest.

In most states, physicians are not required to disclose specific information about themselves [18]. In *Johnson v. Kokemoor*, however, the court held that a physician may have a legal duty to disclose his or her level of experience with a given technique when a reasonable person would expect to be told this information. The case arose after a patient suffered complications from an aneurysm clip procedure performed by a physician whose lack of experience she was unaware of. The experience of the physician was viewed as a piece of information that was material to an informed decision about the procedure [19].

Given that requirements for informed consent are relatively vague and undefined and the exceptions are few, it is in the physician's best interest to inform patients thoroughly about proposed treatment options, ascertain that they understand their choices, and secure their consent. Doing so will help provide quality patient care and avoid exposure to legal action.

References

1. *Schloendorff v Society of New York Hospital*, 211 NY 125, 105 NE 92 (Ny Ct App 1914).
2. *DiFilippo v Preston*, 53 Del 539, 173 A2d 333 (1961).
3. *Canterbury v Spence*, 464 F2d 772 (DC Cir 1972).
4. *Barcai v Betwee*, 50 P3d 946 (Haw 2002).
5. *Gates v Jenson*, 595 P2d 919 (Wash 1979).
6. *Moore v Regents of University of California*, 51 Cal3d 120, 793 P2d 479 (1990).
7. *Jandre v Physicians Insurance Co of Wisconsin*, 330 Wis 2d 50, 792 NW2d 558 (Wis Ct App 2010).
8. *Arato v Avedon*, 5 Cal 4th 1172, 858 P2d 598 (1993).
9. *Nixdorf v Hicken*, 612 P2d 348 (Utah 1980).
10. *Truman v Thomas*, 27 Cal 3d, 611 P2d 902 (1980).
11. *Mallett v Pirkey*, 171 Colo 271, 466 P2d 466 (1970).
12. *Gilmartin v Weinreb*, 324 NJ Super 367, 735 A2d 620 (NJ App Div 1999).
13. *Dunham v Wright*, 423 F2d 940 (3rd Cir 1970).
14. *Stewart-Graves v Vaughn*, 162 Wash 2d 115, 170 P3d 1151 (2007).
15. *Salgo v Leland Stanford Jr Univ Bd of Trustees*, 154 Cal App 2d 560, 317 P2d 170 (Cal Ct App 1957).
16. *Cornfeldt v Tongen*, 262 NW2d 684 (Minn 1977).
17. *Carr v Strode*, 79 Hawai'i 475, 904 P2d 489 (1995).
18. Menikoff J. *Law and Bioethics: An Introduction*. Washington, DC: Georgetown University Press; 2002.
19. *Johnson v Kokemoor*, 545 NW2d 495 (Wis 1996).

Bryan Murray is a third-year law student at the University of Pittsburgh School of Law in Pennsylvania. He is in the school's health law certificate program.

Related in VM

The Legal Boundaries of Informed Consent, August 2008

Discounting a Surgical Risk: Data, Understanding, and Gist, July 2012

A HISTORY AND THEORY OF INFORMED CONSENT

Ruth R. Faden
Tom L. Beauchamp

7

The Concept of Autonomy

At this point we move from historical analysis to conceptual analysis, from the history of informed consent to its nature. This shift requires us to analyze the concept of autonomy and its relationship to informed consent. As the histories in Chapters 3 through 6 indicate, informed consent is rooted in concerns about protecting and enabling autonomous or self-determining choice by patients and subjects. Accordingly, in the present chapter we provide a theory of autonomous action that is adequate to express what is respected or protected by informed consent. In Chapter 8 this analysis of the nature of *autonomy* provides the essential foundation for our analysis of the nature of *informed consent.*

Autonomy and Informed Consent

We begin by introducing two significant distinctions: (a) between autonomous persons and autonomous actions, and (b) between substantially autonomous actions and those that are less than substantially autonomous. The goal of a *substantially autonomous action* will later play a prominent role in our presentation of informed consent.

Distinguishing Persons and Their Actions

Consents and refusals are actions. As we shall see in Chapter 8, informed consents are acts of autonomous authorizing—and, in the case of refusals, of declining to authorize. Somewhat paradoxically, autonomous actions can be performed not only by by *autonomous persons*—those who generally, but not always, act autonomously—but also (on some occasions) by *nonautonomous persons*—those who generally, but not always, fail to act autonomously. Autonomous persons are the kinds of persons who qualify to choose and consent precisely because they are autonomous. The nonautonomous person—for example, a one year old or a severely demented person—is usually not the right type of agent to qualify; he or she is not a person whose choices, consents, or refusals ought to be treated as binding.

Theories of autonomy seem typically to be theories of the autonomous person, as exemplified by the following thoughtful portrait of the autonomous person or "particular personality ideal" drawn by Stanley Benn:

> To be a chooser is not enough for autonomy, for a competent chooser may still be a slave to convention, from his milieu. He assesses situations . . . by norms . . . absorbed unreflectively.
>
> The autonomous man is the one . . . whose life has a consistency that derives from a coherent set of beliefs, values, and principles, by which his actions are governed. . . .
>
> The principles by which the autonomous man governs his life make his decisions consistent and intelligible to him as his own; for they *constitute* the personality he recognizes as the one he has made his own.[1]

Benn's theory features, as he puts it, a "*kind* of agent": The autonomous person is consistent, independent, in command, resistant to control by authorities, and the source of his or her basic values and beliefs. The person's life as a whole expresses self-directedness.

One of the problems with such a theory of the autonomous person is that few choosers, and also few choices, would be autonomous if held to such standards. Benn has correctly emphasized that his theory presents an aspirational *ideal* of autonomy that distinguishes what he calls "normal choosers" from ideal choosers. This status as an ideal theory makes his account too enriched for our purposes. The theory is elevated above what generally is and should be meant in moral and legal theory by expressions such as "respect for autonomy" and "exercise of autonomy," which are broad enough terms in ordinary language and in our theory to include "normal choosers" and their choices. It would follow from an exclusive reliance on Benn's analysis of autonomy that we need not honor and respect *as autonomous* most of the prudent and responsible choices made most of the time by normal choosers. An autonomous choice would presumably also have to conform to an autonomous person's elected life plan—a starkly rigorous requirement—in order to satisfy his ideal standard.[2]

Theories of the autonomous person need not be so demanding, of course. A plausible account of the autonomous person could, for example, be restricted to the enumeration of certain capacities for and patterns of recurrent autonomous action: resistance to social conformity, reflectiveness, understanding and insight, resistance to manipulations attempted by others, and the like. Unfortunately, there is no widespread agreement at present regarding the definitive characteristics of the autonomous person, just as there is no agreement regarding the characteristics of autonomous action.

The distinction between the capacities of persons and their particular actions has been ignored in the literature on informed consent, with unfortunate results that have eventuated in complaints such as the following by William Thompson, in a discussion of tests of patients' abilities to understand disclosed information:

> [General findings] that patients often fail to remember and perhaps fail to understand what they have been told—are interesting and important, but do not resolve the basic question of patients' *capacity* to exercise informed consent. People's failure to understand a *particular disclosure* may result from inadequacies in that disclosure rather than inadequacies *in the patient*. Hence these studies do not tell us whether patients are capable of understanding when properly informed. . . . [Last two italics added.][3]

It is always an open question whether an autonomous person with the capacity to give an informed consent actually has, in any specific instance, given an informed consent, in the sense of making an autonomous choice to authorize or refuse an intervention. The *capacity* to act autonomously is distinct from *acting* autonomously, and possession of the capacity is no guarantee that an autonomous choice has been or will be made. (The capacity to do A may only be *definable* in terms of "A," but the capacity is distinct from all instances of A, just as all dispositions are distinct from their exercise.) An autonomous person who signs a consent form without reading or understanding it is *qualified* to give an informed consent, but has failed to do so. Similarly, if an autonomous person acts in a way he or she did not intend—for example, by accidentally signing form x when he or she meant to sign form y—then the act of signing x is not autonomous, although it is the act of an autonomous person.

It is a common pattern in contexts of informed consent for autonomous persons to fail—for a wide variety of reasons—to give informed consents, even when their acts are formally certified or recognized as informed consents *because* they are acts of autonomous persons. The autonomous person may fail to act autonomously in a specific situation if ill in a hospital, overwhelmed by new information, ignorant, manipulated by a clever presentation of data, and so on.

This helps explain why autonomous actions are examined in this book rather than autonomous persons. Our primary interest is in the nature of informed consent—*what* an informed consent *is*—rather than problems of *from whom* consent should be or may be obtained. Our view is that informed consents and informed refusals are particular actions; the goal of informed consent requirements is to enable patients and subjects to perform these actions, that is, to make substantially autonomous choices about whether to authorize a medical intervention or research involvement.

Degrees of Autonomous Action

Throughout a long history of reflection on autonomy, there has been a recurrent temptation to analyze personal autonomy in terms of two implicit models: (1) a *freedom* model, and (2) an *authenticity* model. The freedom model is employed in dazzlingly diverse ways by writers such as Immanuel Kant, various German and British Idealists, and Isaiah Ber-

lin. The analysis of autonomy in terms of authenticity—"one's own" actions, character, beliefs, and motivation—is developed in no less diverse ways by such writers as Benn, Gerald Dworkin, and various writers in the existentialist tradition. In both models, theories are often dependent on analogies to the autonomy of political states, where "autonomy" has variously referred to popular sovereignty, citizen participation, independent nationhood, nongovernance by alien forces, control by citizens, and the like. However, other theories try to develop an account divorced from a background in political theory.

Our analysis is indebted to these influential theories, but we depart from them in major ways. We analyze autonomous action as follows:

X acts autonomously only if X acts
1. intentionally,
2. with understanding, and
3. without controlling influences.

We believe, but do not here *argue*, that these three conditions are the basic conditions of autonomous action. Whether each condition in 1–3 is a *necessary* condition of autonomous action—as we believe—and whether they are jointly *sufficient* conditions—a more difficult thesis to sustain—are not matters on which we shall speculate in detail, because such claims would need far more argument than can be attempted here. In the above schema we have construed these conditions as necessary but without any mention of their sufficiency, a problem to which we return in the concluding section. Until we reach that section we shall assume only that each of these conditions is necessary.

The first condition, the condition of intentionality, is not a matter of degree: Acts are either intentional (and therefore potentially autonomous) or nonintentional (and therefore nonautonomous). By contrast, the conditions of understanding and lack of controlling influences (which we refer to as the condition of noncontrol, that is, noncontrol by the influences of others) can both be satisfied to a greater or lesser extent. Thus, actions can be autonomous by degrees, as a function of different levels of satisfaction of these two conditions. The conditions of understanding and noncontrol may be placed on a broad continuum from fully present to wholly absent, as represented by the diagram in Figure 1, which uses the language of *substantial* satisfaction of the conditions. For reasons discussed later, we designate this point on the continuum as a watershed for autonomous actions of the sort that qualify as informed consents.

The hypothetical construct of a *fully* autonomous action—represented by the left-most extremes of the two parts of the diagram in Figure 1—graphically depicts the idea of degrees of autonomy. A fully or ideally autonomous action is an intentional action that is fully understood and completely noncontrolled by the influences of others: Lack of under-

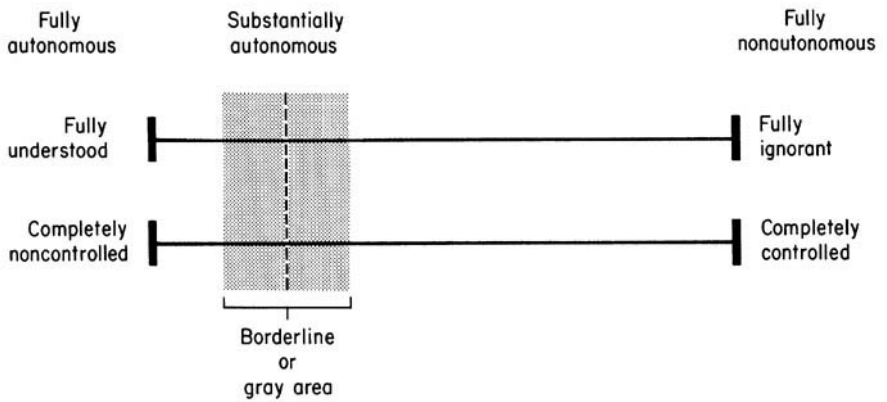

Fig. 1. Degrees of autonomy of intentional actions

standing and control by others do not prevail in any degree. But as the conditions constituting an autonomous action are stripped away by degrees from this hypothetical circumstance, pieces of the interlocking machinery are eroded and the resultant action increasingly becomes less autonomous. With sufficient stripping, the behavior becomes nonautonomous.

The claim that actions are autonomous by degrees, rather than categorically autonomous or nonautonomous, may at first seem a surprising thesis. Yet it is a direct consequence of the conditions that define autonomous action. A person can thoroughly understand an abundance of relevant information at a moment of action or can have only limited understanding. The more information the person understands, the more enhanced are the possibilities for autonomy of action. The form of control exerted by the influence attempts of others also affects the degree to which an action is autonomous. It is noncontroversial, for example, that there are degrees of threats and degrees of abilities to resist threats. Our claim is merely the corollary that some threats render actions more nonautonomous than others.

Another corollary of the degrees thesis, and of our analysis in terms of the three necessary conditions of autonomous action, is that there are *both* different conditions under which *and* different degrees in which autonomous action may be enhanced or compromised. Thus, a person can satisfy one or more of our three conditions, while failing to satisfy another. For example, a person may be fully controlled by a robber's threat of severe harm in yielding his or her money, yet may understand perfectly the nature and implications of the action taken, and may intend to take it. Such an act of compliance would *not*, however, be autonomous, because the condition requiring lack of control by others—like each of the conditions 1–3—is a necessary condition of autonomous

action; and it is not satisfied under the circumstances, not even by degrees. A coerced action is not made any more autonomous if the person's understanding is improved. Thus, if any one of the three conditions is unsatisfied, the act is *non*autonomous. If, however, this is not the case—that is, the act is intentional and there is neither full ignorance nor full control by others, then the degree of autonomy that can be ascribed to an action is correlated with the degrees to which the understanding and noncontrol conditions are fulfilled.

Although this analysis of acts as "autonomous by degrees" is, we believe, conceptually correct, there are also contexts in which it is necessary to establish thresholds above which all acts are *treated as* autonomous and below which all acts are treated as nonautonomous. The decision how and where to draw this line is invariably based on moral and policy considerations, and no sharp line can be drawn purely on conceptual grounds to distinguish autonomous from nonautonomous action. As we shall see, in informed consent contexts we should look neither for *full* autonomy nor merely for *some evidence* of autonomy. The proper threshold is the general goal of *substantially* autonomous action.

Substantially Autonomous Actions

Literature on informed consent overflows with statements that subjects and patients are not qualified to "really understand" that to which they consent.[4] Such statements frequently imply that "full autonomy" is the only instance of real autonomy, and that if a patient or subject cannot exhibit full understanding and full independence from the control of others, then his or her decisions are not informed consents.

Such statements are premised on faulty and unargued assumptions about "real" or "true" autonomy. Like Benn's theory, these assumptions are expressions of *ideal* rather than normal or adequate degrees of autonomy. Such aspirational objectives are worthy ideals, but falling short of them is an inevitable reality. This is true of informed consent, and of virtually all comparably consequential decisions and actions in life. To chain informed consent to *fully or completely* autonomous decisionmaking stacks the deck of the argument and strips informed consent of any meaningful place in the practical world, where people's actions are rarely, if ever, fully autonomous.

As our history of informed consent in Chapters 3 through 6 suggests, consent requirements did not arise with the hope of enabling people to make decisions about health care or research participation that were *more* autonomous than decisions of comparable consequence made in other arenas of life. Much of the initial impetus came from concern that people were *less* able to act autonomously as patients and subjects than elsewhere. Assuming comparable capacity for decisionmaking, one's appreciation of information and independence from controlling influ-

ences in the informed consent setting need not outstrip one's information and independence in making a financial investment, hiring a building contractor, or accepting a job offer. The goal, realistically, is that such consequential decisions are *substantially autonomous*—that is, near the fully autonomous end of the continuum.

In Figure 1, substantially autonomous acts are somewhere between midpoint and fully autonomous, represented symbolically by the dotted line (the middle of the "borderline" area). Any exact placement of this line risks the criticism that it is "arbitrary," because the range for placement is wide, and controversy over any attempt at precise pinpointing is a certainty. Yet the label "arbitrary" is too quick, because the placing of threshold points marking substantially autonomous decisions can be carefully reasoned, if not precisely located, in light of specific goals. We believe that substantial autonomy is a reasonable and achievable goal—an appropriate threshold—for decisionmaking and participation in research or medical interventions, no less and no more than elsewhere in life. The criteria of substantiality for any particular action or type of action will be treated throughout this book as a matter that can only be addressed in a particular context, rather than pinpointed through a general theory. We use the model of substantiality as a standard that patients and subjects might realistically satisfy, and that professionals are morally bound to assist them to satisfy if informed consents should be obtained.

It is a separate issue whether this criterion of substantial autonomy should be the only source for distinguishing consents or refusals that are "valid" from those treated as "invalid." There may be compelling policy or moral justifications in some contexts for adopting consent requirements that establish a threshold below the level of substantial autonomy, in effect treating *less than substantially* autonomous consents as *valid* or, more precisely, *effective* consents. Some of these issues are considered later, but at this point we must turn to the prior matter of analyzing the three conditions in our theory of autonomous action. In Chapters 8 through 10, these will serve, by augmenting the analysis, as conditions of *informed consent*, just as they serve here as the conditions of *autonomous action*.

Three Conditions of Autonomous Action

The Condition of Intentionality

Intentionality is a conceptually necessary condition of autonomous action and in its own right one of the most important concepts in the history of philosophy. Although he never used the word "autonomy," eighteenth-century philosopher Thomas Reid had already seen the vast importance of intention for a theory of what counts as *one's action*:

> Every man knows infallibly that what is done by his conscious will and intention, is to be imputed to him . . . and that whatever is done without his will and intention, cannot be imputed to him with truth. . . . If he intended and willed it, it is his action in the judgement of all mankind. But if it was done without his knowledge, or without his will and intention, it is as certain that he did it not. . . . What I never conceived nor willed, I never did.[5]

The Nature of Intentional Action. But what does it mean to act "intentionally," and is Reid's sweeping characterization correct? Clearly a mere *act of intending*, as when we intend to take our medicine but never get around to it, does not qualify as an *intentional action*. We can intend many things without *acting* on such an intention. "Pure intending," as Donald Davidson has called it,[6] can occur without any practical action or consequence other than the act of intending. Only if action occurs with an intention to perform the action is there an intentional action.

Such intentional actions are not events that merely happen to people. As intentional, they cannot be accidents and cannot be performed inadvertently or by mistake. Nor can it be that another agent uses physical force to push the person into "doing" it. Instead, the actor must have the sense that he or she makes the act happen, that he or she is its author or agent. But even under this description, the scope of the concept is too broad and the meaning too inspecific. The psychological and philosophical literatures abound with diverse analyses of intentional action that treat the notion as synonymous or co-extensive with actions variously described as volitional, deliberate, self-regulated, self-controlled, willed, freely-willed, reasoned, goal-directed, planned, or instrumental. To complicate matters further, the law offers a substantial literature on intentional action that addresses related problems such as criminal *mens rea* and tortious intent.[7]

Despite the diffuse character of the literature, certain core themes are widely shared in contemporary treatments of intentional action in philosophy and psychology. Chief among these is that intentional acts require plans: the integration of cognitions into a blueprint for action. An action plan is a map or representation of the strategies and tactics proposed for the execution of an action.[8] The relationship between intentional action and action plans is straightforward. For an act to be intentional, it must correspond to the actor's conception—his or her plan—of the act in question (although a planned outcome might not materialize).

Imagine, for example, that Toi intends to chop an onion. Toi's plan for chopping an onion entails executing a highly routinized sequence of actions, chief of which is a series of peeling and chopping movements. Note that Toi need not have deliberated about or reflected upon her onion-chopping plan—an actor need not be consciously aware of all or even any of the behavioral steps in the plan, although if there is a devia-

tion or slip in the execution of the plan, that difference may come to the actor's attention.[9] Now imagine that on the way to the cutting board the onion falls from Toi's hand into a working food processor, and the onion is chopped. It might in some ironic sense be said that Toi "chopped" the onion, yet she did not do so intentionally, even though she wanted and intended to chop the onion. Whatever action or behavior was performed, Toi was aware that it did not correspond to *the way* she intended—planned—to do it. Her "chopping" of the onion was, as it turns out, accidental.

This example illustrates how an actor's intention in the action of doing X includes a conception of how X is to be done. Whether a given act, X, is intentional, depends on whether in performing X the actor could, upon reflection, say, "I did X as I planned," and in that sense, "I did the 'X' I intended to do." This is true of even basic actions like moving one's arm, as contrasted with accidental movements such as having one's arm jolted.

In our account of intentional action, the importance of the contrast between an unplanned doing of something and an intentional doing is paramount. Things that agents do by accident are but one subclass of the larger class of unintentional doings. Other candidates include things that persons do inadvertently, certain habitual behaviors, and instances of so-called occurrent coercion in which a person is physically forced by another to do something. However, the most prevalent and relevant kind of unintentional doing—and the kind that offers the most intuitive and informative contrast to intentional action—is the doing of something by accident.

In a heavily condensed formulation, we can say that intentional action is action *willed in accordance with a plan*, whether the act is wanted or not. Consider the following case of accidental behavior, which is not intentional because it fails to satisfy this condition: A man who intends to follow a physician's prescribed dosage opens a medicine cabinet, takes out a prescription bottle, and devours four pills. By accident, the man takes a fatal overdose. It was not his intention to deviate from doctor's orders and certainly not to kill himself. He did both, but he did not do them intentionally, even though he acted intentionally in devouring four pills.

Many questions remain about the cognitive events or processes that are required for intentional action, as do questions about the conditions under which—both phylogenetically across species and developmentally within species—the kind of mental life necessary for intentional action emerges.[10] Fortunately, such issues about the details of intentional action can be sidestepped here, because the relevant problems can all be handled, for our purposes, under the discussion of the condition of understanding. There is in general a close correspondence between understanding that one is doing "X" and doing "X" intention-

ally.[11] In Chapter 9 we argue that in informed consent contexts the association between intentionality and understanding is even closer than in many other contexts of autonomous action. This view has led us to be briefer in this volume in our examination of intentionality than in our analysis of the other two conditions of autonomous action (which are examined in Chapters 9 and 10).

Wanting, Willing, and Acting Intentionally. For a theory of informed consent, a central problem about intentionality is whether special kinds of wants or desires are necessary conditions of intentional action. It could be argued that foreseen acts the actor does not want or desire to perform are not intentional. Alvin Goldman uses the following example in an attempt to prove that such actions are unintentional:[12] Imagine that he, Goldman, is taking a driver's test. The desired act is that of convincing the examiner that he is a competent driver. While driving, Goldman comes to an intersection that requires a right turn. He extends his arm to signal for a turn even though he knows it is raining and thus that, in signaling for the turn, his hand will get wet. According to Goldman, his signaling for a turn is an intentional act. It is a means to achieving his goal of convincing the examiner that he is a competent driver. By contrast, his getting a wet hand is a *nonintentional* act, a mere "incidental by-product" of his plan of action. (For this reason, some philosophers prefer to say that it was an event or occurrence, not an act at all.[13])

In our view, getting the hand wet *is* intentional. It is willed in accordance with a plan in the minimal sense of not being accidental, inadvertent, habitual, and the like, and this fact is for us (unlike Goldman) sufficient for saying that it is intentional.[14] There is nothing about the nature of intentional acts, as opposed to accidental occurrences, that rules out aversions or what one desires to avoid. In this case, Goldman's motivation for getting his hand wet may reflect *conflicting* wants and desires, and in the most straightforward sense of "want," Goldman does not want a wet hand; but this does not render the act of getting the hand wet less autonomous.[15] As Reid suggests, it is something the person did with his "conscious will and intention." Goldman could say to his wife that in sticking his arm out the window he did not act intentionally to shrink his shirt, which was an accident, but he cannot say with credibility that it was an accident that his now ruined shirt got wet. In short, causing the shrinking of the shirt was accidental, but not the getting of the shirt wet.

Goldman's analysis raises profound questions about the individuation of actions that we cannot hope to resolve here. Whether anticipated but undesired events such as wet arms obtained in signaling turns should be considered (1) separate acts ("the hand's getting wet" being a separate event from "the signaling for a turn"), (2) different ways of describing the intended act, or (3) simply a consequence of the intended act is a thorny philosophical problem that we will by-pass through our aforestated thesis that what the person does intentionally must have been

willed in accordance with a plan, although not necessarily wanted. Let us now explore the claim that an intentional action is an action willed in accordance with a plan, even though it may not be wanted.

Some intentional acts are wanted or desired for their own sake, and not for the sake of something else. For example, a lover of swimming may want to swim for the sake of the experience of swimming and not, as may be the case for another person, for the sake of a tanned or fit body. Let us call wanting something for its own sake *intrinsic* wanting. Other intentional acts are wanted not for their own sake but for the sake of something else. In such cases, the actor believes them to be the means to a goal wanted in the primary sense. Let us call wanting for the sake of something else *instrumental* wanting. Only if at least one of these two forms of wanting is involved would Goldman consider an act intentional.

However, other intentional acts are not wanted in either of these two senses. An actor may view these acts as, *ceteris paribus*, altogether *undesirable* or *unwanted*. They are performed only because they are entailed in the doing of other acts that are wanted. One could say that such acts are foreseen, "wanted acts" in the sense that the actor wants to perform them in the circumstances more than he or she wants not to perform them. But it is better to overthrow the language of wanting altogether and to say, as Hector Castañeda puts it, that they are *tolerated* acts.[16] They are not so undesirable as to cause the actor to choose against performing the desired act that entails them. In such a situation, both kinds of acts—the desired act and the otherwise undesired act—are, as we analyze action, *intentional*. Let us call such intentional actions ones of "tolerating" rather than "wanting."

Consider the following three examples, which illustrate the three forms of intentional action, two forms of wanting and one of tolerating: (1) John has always liked facial scars on men. He particularly likes sharp, jagged scars over the cheekbone, although, as with many aesthetic preferences, John is unable to explain why he likes them. John decides he wants just such a scar on his own face and enlists the services of a plastic surgeon. Here John intrinsically wants the scar; he intentionally consents to *surgery*, and he intentionally consents to being *scarred* by surgery. (2) Harry is involved in a personal injury suit in which he is trying to recover for lacerations to his face and neck suffered in an automobile accident. Harry is counting on the money in the settlement to reduce some outstanding gambling debts; he worries that the scar on his face is healing too well. He decides to enlist the services of a plastic surgeon to freshen the wound to make it jagged and sharp. Here, Harry wants the scar instrumentally; he intentionally consents to *surgery* and he intentionally consents to *being scarred* by surgery. (3) Don has skin cancer on his right cheek. He is eager to have surgery until his physician informs him that, despite use of the best technique, he will be left with a sharp, jagged scar. Don is horrified at the prospect of disfigurement and desperately

wants to avoid being scarred. Nevertheless, after becoming convinced that there is no meaningful way that he could consent to surgery but refuse the scarring, and after considering the alternative of refusing surgery, Don intentionally consents to *surgery*, and in so doing he intentionally consents to *being scarred* by surgery. Although his consenting to a facial scar is a tolerating and not a wanting of the scarring, the intentional act of consenting to the scarring is no less Don's *own* act than is his consenting to surgery. Don, then, intentionally consents to being scarred by surgery.

This last sentence has an intuitive ring of oddity. It seems strange to insist that Don intends to be scarred, just as Harry and John intend to be scarred. This oddity has led Goldman, John Searle, and others to interpret intentional action more narrowly than we have.[17] The model of wanting, as we may call it, is the only model that applies to intentional actions for them, so that tolerated actions cannot be intended. By its very presuppositions, such analysis cannot accommodate, *as intentional*, tolerated acts such as getting one's hand wet and consenting to scarring. By contrast, we have interpreted intentional actions more broadly to include any action willed in accordance with a plan, including tolerated as well as wanted acts.

We are thus using a *model of willing* rather than a *model of wanting* to capture the breadth of intentional action. In Don's case, for example, while it may seem odd to say that he intentionally consented to being scarred by surgery, it seems still more odd to say that Don's consent to being scarred was unintentional or was a nonintentional act, as Goldman prefers to say. For those who find both a model of willing and a model of wanting awkward, the following reformulation may prove more appealing: Don intentionally consents to *surgery* that he knows will cause him to have a facial scar. The important matter is that a person's willing acceptance of tolerated outcomes is properly describable as intentional.[18]

Still, there is something unsettling and odd about this conclusion. Don's intention in undergoing surgery does not seem to be that of being scarred; indeed, the scarring is a reason against undergoing the surgery. We seem to be saying that Don intends to be scarred and intends not to be scarred. It seems self-contradictory to say that Don intends to minimize the scarring and at the same time say that he intends the scarring. From the physician's perspective, it appears that the doctor intends to scar Don but will do his best not to scar him, almost as if he intends not to carry out his intention. Can the doctor hope not to do what he intends to do? Can Don?

The way out of these problems is the following: While we do treat "willing" and "intending" as synonymous, it is not our claim that "being willing" is synonymous with "acting intentionally." Acting under the condition of being willing is only one kind of intentional or willed action.

In other cases, one is not merely willing to do something but positively eager to do it. Whether done from a condition of being willing or being eager is irrelevant to its being intentional; the act is intentional in both cases. Thus, "Don intends to be scarred" can be read as "Don is willing to be scarred" in this case; and "Don intends to minimize the scarring" can be read as "Don is eager to see that the scarring be kept to a minimum." Don acts intentionally in authorizing the surgery and intends, quite consistently, both of the above.

Acts that can be described as intentional consents involving acts of toleration are common in informed consent settings. Undesirable consequences or risks of harm that attend particular procedures generally fall into this category. Case law hinges almost entirely on whether disclosures about undesirable outcomes or risks have been made, and, if not, whether they represent tolerable outcomes under the circumstances or whether knowledge of them would have altered the patient's intentions and behavior. Consider, for example, the case of *Bang* v. *Charles T. Miller Hospital*: Mr. Bang did not intend to consent to sterilization, which was an outcome of prostate surgery, a surgery to which he *did* in fact give his formal, intentional consent.[19] (Under negligence theory the physician-plaintiff would not have prevailed if it could have been demonstrated that sterilization is a *tolerable*—even if undesirable—outcome of prostate surgery, but only if such an outcome would not cause a reasonable person to refuse consent.) If Mr. Bang had consented with an understanding of the inevitability of sterilization, then the infertility would have been a tolerable outcome and his consent to that outcome would have been an intentional act of toleration.

Are There Degrees of Intentional Action? Because our theory of autonomy is based on the concept of degrees of autonomy, the following question might be raised: Are all intentional acts equal in degree of intentionality? That some acts are *more intended* or *more intentional* has an intuitive appeal, yet it is impossible, we believe, to support this intuition by a theory. It is unclear what (if anything) would make actions more or less intentional, although there are two not entirely implausible grounds for explicating the idea of degrees of intentional action. These are *mindfulness of willing* and *correspondence with an action plan.* According to the first—mindfulness of willing—acts are more or less intentional, according to how emphatically the will is at work in expressing them. Less intentional actions are in some degree automatically executed, with little if any cognitive awareness. More intentional actions, by contrast, require what Miller, Galanter, and Pribram call emphatic inner speech, a real "effort of the will."[20] This effort may involve deliberate planning, careful monitoring, perhaps even written notes for the act to be successfully conceived and executed. For example, when a person is first learning to drive a car, every act of learning to control and react requires concentration, deliberation, and active internal awareness. However,

once the person has become an accomplished driver, the same acts of steering, engaging the passing lights, and the like can be performed automatically—in effect, mindlessly.

A second possible way of explicating the notion of degrees of intention is through the degree of correspondence between an action and an action plan that prefigures *what* the actor intends to do and *how* he or she intends to do it. The more closely the action corresponds to the action plan, one might argue, the more it can be said that the action the person did perform was the action the person intended to perform. (One can, of course, ditch or modify an action plan in mid-stream, in which case a new action plan is the test model for assessing correspondence.) This degree of correspondence between the action performed and the action plan depends not only on the sheer number of discrepancies from the plan—a quantitative consideration—but also on the importance of each discrepancy—a qualitative consideration. An action can deviate from an action plan in trivial ways that have no significant effect on the action's character, but an action can also deviate from an action plan in major ways, so that the act is almost unrecognizable as the act intended under the action plan.

However intuitively attractive the idea of degrees of intentional action may be, it is a theory to be avoided. Acts are either intentional or nonintentional, not somehow on a continuum. As Jay Wallace has pointed out to us, one can say, perfectly correctly, that actions can be more or less *deliberate* without having to say that they are more or less *intentional.* The extent to which an intentional act is deliberate may well be measured by criteria of consciousness, reflectiveness, and correspondence to an action plan. But even comparatively nondeliberate acts, such as automatically breaking or flipping a light switch, may still be intentional in exactly the same respects as deliberate actions like consenting to surgery. All things considered, then, it seems best to avoid succumbing to the tempting idea of degrees of intentional action.

It is unlikely, in any event, that consents and refusals would fail of substantial autonomy because they fail the two grounds just discussed. Acts that did fail these grounds would also likely fail to satisfy the condition of understanding, a subject to which we now turn.

The Condition of Understanding

An action cannot be autonomous if the actor fails to have an understanding of his or her action. This condition is of special importance for a theory of informed consent. Clinical experience suggests that patients and subjects exhibit wide variability in their understanding of information about diagnoses, procedures, risks, prognoses, and the like. Although some patients and subjects are calm and attentive, others are nervous, anxious, or distracted in ways that affect understanding. The quality of

autonomous decisionmaking may therefore differ dramatically, in accordance with how well a person understands.

But what does it mean for a person to understand? There is no consensus answer or approach to this question in philosophy or psychology, where there is a surprisingly brief history of discussion.[21] Some of the great classic works in epistemology, including Locke's *An Essay Concerning Human Understanding* and Hume's *An Enquiry Concerning Human Understanding*, are by title treatises on "the understanding"; but these works are more focused on ideas, belief, perception, mental concepts, processes of knowing, and the like than on the analysis of understanding. "Understanding" primarily refers in these writings to the intellect or to faculties of knowledge and judgment. Although many contributions in past and present epistemology are directly pertinent to the analysis of understanding, we endorse a quip by Arthur Danto: "Think how much we understand in comparison with how dark the concept of understanding remains to this day!"[22]

It is questionable whether in psychology the understanding of understanding is more advanced than in philosophy. The central question in psychology has been *how* people understand or comprehend human communications, with the emphasis on cognitive and neurophysiological processes. Major discussions have, for example, focused on the structures of memory and knowledge involved in word recognition, sentential understanding, and discourse interpretation,[23] as well as on the effect of stress and neuropsychological factors on learning. Work on these and other topics has relied on operational definitions of key terms, in almost total isolation from conceptual and comprehensive experimental treatments of the nature of understanding—although, as we shall see in Chapter 9, some of this work is relevant to problems in informed consent that pertain to the understanding of disclosed information.

Some writers have a ready explanation for why psychology has failed to advance analysis of the concept of understanding. They surmise that this task is primarily philosophical; to refer to a "theory of understanding," without being careful about which *kind* of theory, is to conjure up false expectations and confusions. From their perspective, philosophy is the discipline that provides analysis of the *meanings of terms* such as "understand"; psychology is the field that provides a theory of learning about *how we come to understand.* We reject this dichotomy in our examination of understanding. We do not seek simply an analysis of the word "understand," nor an explanation of how empirically we come to understand. Rather, we treat what it is to understand—the demands that must be satisfied in order for it to be true to say that someone understands something.

The Uses of "Understand." Numerous problems must be faced before any such analysis can be successful. "Understand" is a word with several uses, as the following sentences illustrate: "I understand his motives,"

"I understand English," "I understand that the sun is at the center of the universe," "I understand the history of the period," "I understand the word 'leopard'," "I understand the truths of mysticism," "I understand how to pray," "I understand that you have been faithful," "I understand you completely," and "I understand what you are saying."

In the sentence, "I understand how to pray," "understand" entails "competence," as we analyze that term in Chapter 8. Unlike most of the uses in the sentences in the above list, to understand in this sense is to have a practical competence or know-how, and the word is used to express one's possession of it. This can be called "understanding *how* (to do something)," which is equivalent to knowing how.

By contrast, there is a second use of "understand": Sentences like "I understand that you have been faithful" and "I understand that the sun is at the center of the universe" are straightforward propositional knowledge claims. This can be called "understanding *that* (some proposition is true)." Here the analysis of understanding is reducible to the analysis of knowledge—for example, knowledge as justified true belief (or some variant of this analysis). One form of understanding of relevance to this volume is understanding that "I am being asked to consent to this intervention." This is one kind of "understanding that."

Other acts of understanding have to do with human communications, and this takes us to a third use of "understand." Here one does not have to *believe* information in order to understand it. The sentences "I understand what you are saying" and "I understand you completely" do not imply that one believes that what is said is true, only that one apprehends what has been said. This can be called "understanding *what* (has been asserted)." "Understanding what" is correctly analyzable as "understanding that" if the understanding is *that* "proposition p has been asserted." This fact we can acknowledge, only to set it aside. The important point for our purpose—the analysis of autonomy and informed consent—is that the typical pattern of understanding in informed consent settings is for patients or subjects to come to understand *that* they must consent to or refuse a particular proposal by understanding *what* is communicated in an informational exchange with a professional.

Although this pattern of communication is typical of contexts of informed consent, the full pattern is not essential for understanding. As we shall see in Chapter 8, a consent can be based on substantial understanding quite independent of any professional exchange or disclosure. For both our theory of autonomous action and our theory of informed consent, the kind of understanding at issue is a special case of propositional understanding or "understanding that." Specifically, it is a person's understanding that his or her action is an action of a certain description with consequences of a certain description. Thus, we need an account that can identify the conditions under which a person understands the nature and implications of his or her actions. More colloquially, we need a way of analyzing the question, so frequently asked in

both psychiatric and informed consent situations, "Do you understand what you are doing?" This question is usually translatable as, "Do you have justified beliefs about the consequences of what you are doing?"

Understanding One's Action. What does it mean to understand that one is performing a certain act? Thus far, this kind of understanding has been treated in the context of informed consent largely through law and psychiatry, which appeal to criteria governing competence and insanity determinations. Understanding has been given a narrow compass in these fields that is unsuitable for our purposes. On the other hand, the problems confronting any complete analysis of this kind of understanding are vast. In this volume, we offer but a modest proposal by comparison to that vast literature. Our analysis is intended to convey only the distinguishing characteristics of understanding one's action.

Starting with the polar extreme of *full* or *complete* understanding, let us say, as a tentative hypothesis, that a person has such total understanding of an action if the person correctly apprehends all the propositions or statements that correctly describe (1) the nature of the action, and (2) the foreseeable consequences and possible outcomes that might follow as a result of performing and not performing the action. So interpreted, *full* understanding does not amount to *omniscience*, because this criterion demands only *foreseeable* present and future events, not all such events. However, there are two dangers here, both worth careful attention.

First, it might be argued that our definition is *not demanding enough*: In one perfectly reasonable sense of "understand" we do demand omniscience, or at least more than mere foreseeability. For example, if workers exposed to carcinogens in the workplace have no reason to believe that the chemicals they handle are dangerous, and know everything about the chemicals that is known by anyone at the time, then they know the foreseeable consequences and yet do not understand what they are doing to themselves in handling the materials. Similarly, both patients and surgeons who know all there is to know about the consequences of a surgical procedure in light of available information know the foreseeable consequences, but if there are unforeseen, unknowable consequences that would count as entirely compelling reasons not to perform the surgery, then there is an important respect in which patients and surgeons do not understand what they are doing. Our definition is clearly not broad enough to capture this sense of "understand," especially as an account of *full* or *complete* understanding.

Our way of handling this first problem is to set it aside. The restricted sense of "understand" we shall operate with is tied to an *evidentiary standard* of what it is to understand—a justified belief standard rather than a justified *true*-and-*full* belief standard. We shall return to this point momentarily.

Second, the reverse kind of objection might be made against our definition, namely that requiring apprehension of *all* correct and foresee-

able descriptions of an action and its possible consequences is *too demanding* a standard, one far too removed from any ordinary or conventional understanding of human actions. A person's apprehension of many descriptions of this sort would make no contribution whatever to the person's understanding of the act, at least not in any sense relevant to the personal or social meaning(s) of the action. For example, most human actions involve bodily movements; act descriptions will be available of chemical and atomic processes attending them. But these descriptions may be entirely irrelevant to understanding the action. Therefore, it seems that in analyzing *full or ideal* understanding, the tentative characterization we have given is too broad.

We accept this second caution and so would reformulate the definition as follows: A person has a *full* or *complete* understanding of an action if there is a fully *adequate* apprehension of all the *relevant* propositions or statements (those that contribute in any way to obtaining an appreciation of the situation) that correctly describe (1) the nature of the action, and (2) the foreseeable consequences and possible outcomes that might follow as a result of performing and not performing the action. To the extent that this ideal is less than satisfied, an action is based on less than *full* understanding, and thus is less than a fully autonomous action.

Several problems attend any attempt to evaluate the extent to which a person's understanding of an action deviates from this ideal of full understanding. First, a list or count of the *number* of relevant act-descriptions the person fails to understand will not necessarily express the actual extent of the person's understanding of his or her action. In our discussion of intentional action we noted that not all discrepancies between action plans and actual actions are of equal importance. This is also true of act-descriptions. Some act-descriptions are irrelevant, trivial, and others vital. Even if a person understands many descriptions of his or her act, the failure to understand just *one* description, if it is important enough, can substantially compromise understanding. For example, in the aforementioned real-life case of Mr. Bang, the adequacy of his understanding of his act of consent apparently turned on a single act-description, namely, that his consent involved a consent to sterilization. In Chapter 9 we propose, as a partial explication of the threshold of substantial understanding, a criterion in the informed consent context for distinguishing relatively important act-descriptions from relatively unimportant ones.

How to make this distinction between trivial and important act-descriptions is only a piece of the larger problem of understanding one's action. Another unsettled problem concerns what it means for someone to apprehend or appreciate an act-description *adequately*, whether it is trivial or important. For example, many patients who consent to major surgery understand that as a consequence of their consent they will suffer postoperative pain. Many "understand" that they will experience the pain, but their projected expectations of the pain may be seriously inad-

equate.[24] In some situations patients simply cannot, in advance, adequately appreciate the real nature of the pain, just as some of the students in the Zimbardo mock-prison study discussed in Chapter 5 could not have appreciated in advance the actual intensity of the stress that would be involved in the experiment. Even with prior personal experience it is often difficult to appreciate or anticipate the depth of projected stress and pain.

In such situations there is a respect in which patients and subjects *correctly apprehend* the act-description, "If I consent to this procedure, I will have severe pain." But there is also a respect in which their apprehension or understanding of the act-description is *less than correct*; this deficiency occurs to the extent that their conception inadequately reflects what pain, stress, and the like *are*, how they feel, how they take over the person's responses, and so on. This problem is compounded by the inevitable variation in tolerance of pain and stress, both between individuals and within the same individual over time.

Ignorance and False Beliefs. Other problems for our account of what it means to understand one's action are posed by the issue of false beliefs. Consider John Stuart Mill's much discussed example, in which a man wants to cross what he believes to be an *intact* bridge.[25] Unfortunately, this belief in intactness is false. The bridge is out. Assuming the state of the bridge is *foreseeable*, the man's understanding of the action of walking across the bridge is less than complete because it fails to include at least one relevant act-description, namely, that if he attempts to cross the bridge he will fall in the river.

The relationship of demonstrably false beliefs to autonomous action is relatively straightforward. To the extent that a person's understanding of a proposed action is based on false beliefs about that which would otherwise be relevant to an understanding of the action, performance of that action is less than fully autonomous (even if he or she is "responsible" for the false beliefs on which he or she acts). Of course, some false beliefs are more important as descriptions of the action than others, and these must be given more weight in an assessment of their impact on autonomy. We must also take account of the probabilities involved in belief. If a person believes that a medical treatment will be effective in a particular case, it does not turn out to be the case that the person had a false belief about the general efficacy of the treatment merely because it proves to be ineffective in that case. If the person did not understand the inherently probabilistic character of the judgment of effectiveness—believing, for example, that the treatment always works—then the person would have proceeded on a false belief about efficacy. But we would not want to say that no action is autonomous if the actor is surprised by the outcome.

Even a superficial discussion of the roles of true and false propositions as bases of action surpasses our objectives. But we must consider one problem: Little analytic probing is required to show that it is often dif-

ficult to provide *evidence* of truth or falsity or to achieve *intersubjective agreement* about the truth or falsity of beliefs, and that these difficulties present major problems for informed consent. Consider the following three act-descriptions: (1) "If I consent to this blood transfusion I will burn in hell." (2) "I know that 50% of the patients with my problem do not survive this type of surgery. I am a fighter, and if I consent to this surgery, I *will* survive." (3) "Nothing good can come of my consenting to this procedure, because no matter what new skills and coping styles I might develop, I will never want to live as a quadriplegic." Each of these beliefs could be central and even decisive for a person's decision to consent to or refuse medical treatment. Yet they may be viewed by others, including those seeking consent, as highly questionable, poorly reasoned, or patently false.

All other things being equal, does the holding of such beliefs diminish understanding and thus autonomy? There is no answer to this question that will be free of heavy qualification. One approach to the probabilities and inherent uncertainties that attend many beliefs is to adopt an *evidential* standard for evaluating the acceptability of a belief statement. Rather than invoking a standard of truth or correctness, this approach asks whether, under the circumstances, the person holding the belief is *justified in believing* that it is true—and therefore, if appropriate, *warranted in asserting* to another *that* it is true. This justified-belief standard captures the common-sense conception of reasonable (even if not true) belief and assertion that underlies ordinary social agreements about what is veridical.

There is, of course, more than one standard of evidence, and significant epistemological problems surround what counts as evidence and therefore as justified belief. All evidence gathering must be done within the context of some theory that governs what counts as the facts or as evidence. No evidence is ever independent of the theory that it presupposes, and two or more theories may put forward competing standards of evidence. Yet the criteria for theory-assessment may themselves be in dispute, and there may be no higher-level rules for assessing the criteria in the competing theories.

A justified belief standard provides little assistance, then, in making judgments about the acceptability of particular beliefs if there is little or no agreement on the criteria for determining the *justifiability* of such beliefs. If the justifiability of a belief or the warrant for its assertion is inherently and unavoidably contestable, there may be no adequate grounds for determining whether a given belief compromises understanding, and whether it prevents the person from acting autonomously. While this is a major problem in epistemology of indisputable relevance to discussions of informed consent, it does not provide a sufficient reason to preclude adoption of a justified belief standard for assessing the quality of a person's *understanding*. This is the standard that we shall hereafter adopt.

Accurate Interpretation and Effective Communication. As noted earlier, understanding one's action—a form of "understanding *that*"—is often heavily dependent in informed consent contexts on "understanding what"—for example, understanding what is being disclosed by a professional seeking consent. If a person claims to understand *that* something, X, is the case, evidence for the truth of X must be shown. But in claiming to understand *what* someone said in asserting the truth of X, one must only show that communication of X has occurred, not that what has been communicated, X, is true. In order to understand in a circumstance of communication, one's formulation or representation of what is said must be substantially *accurate* but not necessarily *believed.* If what is communicated and understood is false, that fact too is irrelevant, because it is irrelevant whether one believes it or is justified in believing it. The only vital condition of understanding here can be stated in two words: *accurate interpretation.* Usually such accurate interpretation occurs in the form of effective communication, which is generally analyzable in terms of having justified beliefs about others' statements and intentions. We shall see when we study understanding in Chapter 9 that in many consent situations effective communication is *without peer as the most important form of understanding.*

"Understanding," explicated as accurate interpretation, resembles some analyses of the German word "Verstehen" (to understand), which has received intense attention in writings on the philosophy of history in discussions of the understanding of the actions and motives of other persons. One understands in this sense if and only if one can penetrate and accurately interpret the past inner functioning of another's mind, reliving the person's experiences and rethinking the person's thoughts by a process of inference. As Patrick Gardiner puts it, the claim is that understanding consists in taking a "psychological X-ray photograph," or reconstructing what must have occurred, just as a master detective reconstructs the moves and motives of a criminal.[26]

Although sometimes excessively elevated in importance in German philosophy, this verstehen theory provides a useful insight. To understand what someone has said is often to establish a *correspondence* between one's interpretation or representation of the statement and what the person meant to say. To the extent that such an assumed correspondence is inaccurate, the hearer does not accurately interpret or understand (what has been said). How to explicate the nature of this correspondence or accurate interpretation has proved elusive, to say the least, eventuating in metaphorical appeals to what it might mean for a language or a mind to "picture" a state of affairs. This problem exceeds the range of our discussion. We can nonetheless take the ideas of correspondence, accurate interpretation, and effective communication to be basic. They should provide an adequate basis for our understanding of understanding here and in the expanded discussion in Chapter 9.

The Condition of Noncontrol

A fundamental condition of personal autonomy is that actions, like the actions of autonomous states, are free of—that is, independent of, not governed by—controls on the person, especially controls presented by others that rob the person of self-directedness. The close connection between this condition and autonomy is semantically obvious in that autonomy, self-governance, and self-determination are all treated in dictionaries as synonymous with independence from control by others. As the third condition of autonomous action in our theory, *noncontrol* has equal standing with intentionality and understanding. However, it differs from these other two conditions by not being measured as a positive presence, but in terms of the negative condition of not being controlled by others. "Noncontrol," then, entails that there are not external controls on the action; this independence from control is its relevance for a theory of autonomous action.

In analyzing the condition of noncontrol, we use *influence* and *resistance* to influence as basic concepts in the analysis. Control is exerted through influences in our analysis, but not all influences are controlling. Many influences are resistible, and some are even trivial in their impact on autonomy. Influences come from many sources, and take many forms: Threats of physical harm, promises of love and affection, economic incentives, reasoned argument, lies, and appeals to emotional weaknesses are all influences. They can vary dramatically in degree of influence actually exerted.

These problems of control are often addressed in the literature on informed consent through the concept of voluntariness, which is sometimes treated as synonymous with autonomy. The role and importance of "voluntariness" for informed consent is emphatic and unconditional in the lead principle of the Nuremberg Code (see pp. 153–158):

> 1. The voluntary consent of the human subject is absolutely essential. This means that the person involved should have legal capacity to consent; 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, overreaching, or other ulterior form of constraint or coercion. . . .

Literature on informed consent tends to treat acting *voluntarily* (or voluntary action) as synonymous with acting *freely* (or free action). However, caution is in order about these terms, and likewise about the related notions of acting *intentionally* (intentional action) and acting *willfully* (willed action). We use "willing" and "intending" synonymously, in accordance with our previous analysis of acting intentionally.[27] Both should be sharply distinguished from "control" and "noncontrol": Nonintentional actions are not necessarily controlled by the influences of others; and an *act* can be completely noncontrolled by the influences of others and still be nonintentional. For example, if a person accidentally

signs the wrong consent form, this is an unintentional act of signing (some prefer to say that it is an act of signing mistakenly), but not necessarily an act controlled by others.

Conversely, if a person's actions are entirely controlled by others, these actions may still be intentional, as when a victim at gunpoint intentionally hands over possessions to a robber. A person coerced by such a threat is obviously controlled; the threat deprives the person of freedom, normal opportunities, and voluntary choice—to invoke some of the more time-honored terminologies. Although the coerced person is deprived of the ability to act *autonomously*, the person is not thereby deprived of the ability to act *intentionally*. That is, the person may respond to the threat with an intentional action, albeit one completely controlled by the influence of another person. We can say, without conceptual confusion, that a person who is entirely controlled in performing an act may nevertheless *act intentionally* in performing it.

In some literatures acting voluntarily and willing or intending are taken as near synonyms, whereas in our theory acting intentionally is distinguished sharply from acting in the absence of controlling influences. The word "voluntariness" is sometimes defined so broadly as to be virtually synonymous with both autonomy and freedom of the person. Joel Feinberg, for example, analyzes voluntariness and voluntary choice in terms of the absence of psychological compulsion, the presence of adequate knowledge, and the absence of external constraints.[28] Were we to adopt Feinberg's analysis, the voluntariness condition would become the parent or generic term for autonomous actions. It would be *the* rather than *a* condition of autonomy. In other treatments, "voluntariness" has been loosely related to ideas about agency, authorship, nonsubjection to authority, privacy, personal command, authenticity, and choice—all commonly associated with autonomy no less than voluntariness.

These many confusing associations surrounding the term "voluntariness" are too much, we believe, to combat successfully through a conceptual analysis that attempts to tidy up its meaning, and hence we avoid the word entirely. We substitute a conception of noncontrol that does not have the history and connotation that burdens the terms "freedom," "voluntariness," and "independence."

The Nature of Control and Noncontrol. We also choose the word "control" because it can be shaped in meaning by an account of the nature of control and influence more readily than can terms such as "constraint," "restraint," "rule," "governance," and other terms that are used to express ways of powerfully affecting behavior. Words like "constrain" and "restrain," for example, imply that force, compulsion, or coercion must be at work, but this analysis is distinct from the one that we shall develop of influence and control. One advantage of coining a new technical term like "noncontrolled" is the relative ease with which it can be defined without obstructive, counterintuitive associations.

What, then, do we mean by the terms "controlled act" and "noncon-

trolled act"? We can begin to address this question by examining the polar extreme of a *completely* or *fully* noncontrolled act. Such acts have either (1) not been the target of an influence attempt, or (2) if they have been the target of an attempt to influence, it was either not successful or it did not deprive the actor in any way of willing what he or she wishes to do or to believe. Thus, a person can be *influenced* without in any way being *controlled* by the influence agent, because the person acts on the basis of what he or she wills rather than on the basis of subjection to the influence agent's will and ends.

By contrast, completely controlled acts are entirely dominated by the will of another; they subject the actor to serve as the means to the other's ends and in no respect to serve the actor's own ends. Person A's action controls an action X of person B if A gets B to do X through irresistible influences that would work even if B, left to his or her own ends, in no way wanted to do X. If a doctor orders a reluctant patient to take a drug, and coerces the patient to compliance, then the patient is under the will of the doctor and is fully controlled. If, by contrast, a physician merely persuades a patient to take a drug that the patient is at first reluctant to take, then the patient wills to take it and is not under the will of the doctor.

There are many in-between cases: For example, suppose the physician has made it clear that he or she will be upset with the patient if the patient does not take the drug, and the patient is intimidated. Although the patient is not convinced that it is the best course to take the medication, which would not be accepted if the physician merely offered it as an option, the patient agrees to take the drug because it appears that acceptance will foster a better relationship with the doctor than could otherwise prevail. Here the patient performs the action to a significant extent on the basis of what is willed to personal ends, but only under a heavy *measure* of control by the physician's role, authority, and indeed prescription. Unlike the first case, the patient does not find it overwhelmingly difficult to resist the physician's proposal, but, unlike the second case, it is nonetheless awkward and difficult to resist this rather "controlling" physician. The physician's recommendation is not *ir*resistible, simply *difficult* to resist.

From Control to Noncontrol. Not all influences are controlling, and not all that are controlling are *equally* controlling. Some facilitate choice and are welcomed by persons on whom they operate. In our analysis, there are three *main* categories or forms of influence: coercion, manipulation, and persuasion (hence the title of Chapter 10). However, as Figure 2 exhibits, manipulation is the only form of influence that rests on a continuum. Coercion and persuasion are not continuum concepts: Coercion is always controlling, but not by degrees; persuasion is never controlling and involves no degree of noncontrol. Manipulation, by contrast, *is* controlling or noncontrolling and admits of degrees. This analysis involves

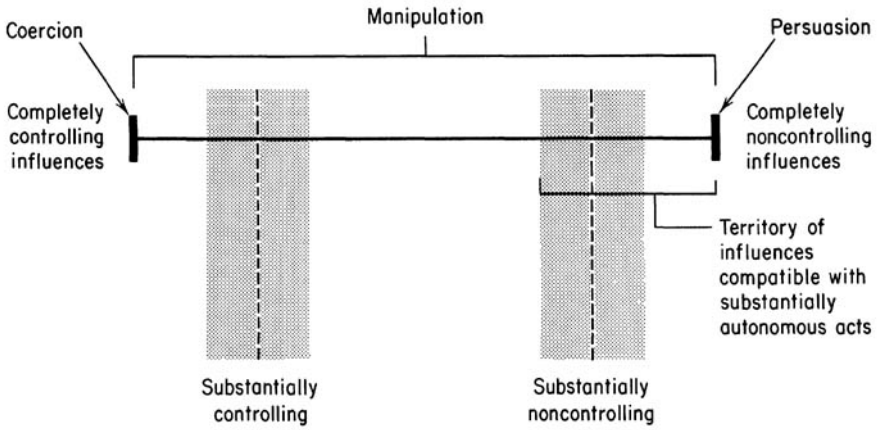

Fig. 2. The continuum of influences from controlling to noncontrolling

some theoretical reform of ordinary conceptions of these notions and so bears further discussion.

Coercion, the most frequently mentioned form of influence in philosophical and legal literature, might be thought to rest in a broad region near one end of the continuum of influences, from influences that wholly control the person to less controlling influences (see Figure 2). But this is an incorrect depiction. Coercive interventions always entirely compromise autonomy by wholly controlling action. At the opposite extreme are forms of noncontrolling influence that do not govern the person, but rather facilitate choice. They can be resisted and it is always up to the person's decision to accept or reject these forms of noncontrolling influence. Persuasion is the most notable example. When persuaded by another, one willingly acts or accepts a belief as one's own. Other forms of influence such as deception, indoctrination, seduction, and the like are forms of manipulation. Some manipulative strategies can be as controlling as coercion or as noncontrolling as persuasion; other manipulations fall somewhere between these endpoints.

This description of a continuum from control to noncontrol that is coextensive with manipulation follows from our earlier discussion of degrees of autonomy. Our primary concern in treating informed consent is with the *extent* to which any form of influence or particular influence renders an action less than substantially noncontrolled and therefore outside the territory of influences compatible with substantially autonomous acts. The dotted lines and shaded areas on Figure 2 represent the ambiguity involved in establishing thresholds for substantial control and noncontrol. There are no definitive criteria for distinguishing with exactitude acts that are substantially noncontrolled from acts that are only slightly more controlled—or for that matter, for distinguishing influ-

ences that leave the person substantially noncontrolled from influences that move toward more control.

Distinctions between the three main categories are not sufficiently refined or refinable to permit all influences to be sharply distinguished and classified as either manipulative or persuasive or coercive. We provide definitions of coercion, manipulation, and persuasion that serve to highlight the distinctions between the categories and to permit proper classification of most cases. However, these definitions are not sufficiently rigorous to withstand all counterexamples presenting "hard cases," and we do not systematically defend our definitions against others that have appeared in the philosophical, legal, and social scientific literatures.

Subjective vs. Objective Interpretations of the Basic Categories. Both objective and subjective interpretations of these categories of influence are possible, but our analysis relies predominantly on a subjective interpretation. Individuals are subjectively influenced in different ways, some persons being far more resistant to particular influence attempts than others. What manipulates or coerces one person may not manipulate or coerce another, even if the same manipulation attempt is directed at both equally. *Being manipulated*, then, is an inherently subjective event, one relative to the person's response. However, the terms "manipulation" and "manipulative acts" are often used, correctly, in a nonsubjective sense. In ordinary language, an act directed at a large audience, such as a television advertisement, may be properly described as a manipulation, or as manipulative, even if it succeeds in manipulating only a small portion of those at whom it is directed.[29]

A fanciful example will exhibit what is at stake in choosing between an objective and a subjective analysis: Imagine that a mischievous daughter wants to go on a date with a boy her mother intensely dislikes. The mother refuses permission. The daughter then produces a caged mouse and says, "Either I go out on the date or I will let the mouse loose in the house." Most mothers, let us suppose, would not be much troubled by the mouse, although terribly upset with the daughter. A typical parental reaction would be to seize the caged mouse and assert dominion over the daughter. However, *this* mother, as the daughter well knows, has a severe, psychologically crippling mouse phobia and will yield to any request when faced with the alternative of a mouse about the house. This mother is *coerced* by the daughter's threat into consenting to the giving of permission for the date, even though the average person (or reasonable person, or rational person, or normal person, or most persons) would not be so coerced in the circumstances.

Because the average (or normal, reasonable, etc.) person would find the mouse threat resistible, the daughter's action is not one of coercion on an *objective interpretation* of the category. The notion of "objectivity" is here tied to average (normal, etc.) reactions (see pp. 32–33 for an analysis of "objective" in this sense). If the average (normal, reason-

able, etc.) person would be able to resist, then on an objective interpretation no threat is coercive, no matter its effect on any *particular* person. This account is unsatisfactory for our purposes, because the subjective impact on the individual is the decisive matter in determining the extent to which the person's action is autonomous. The mother's giving permission for the date *would be coerced* and therefore nonautonomous on our analysis, even if no one else in the world would be successfully threatened in this circumstance by the presence of an uncloistered mouse.[30]

In Chapter 10 we shall explore further the relevance of this distinction between objective and subjective interpretations, as they impact on our understanding of informed consent.

Definitions of the Three Forms of Influence. Our basic definitions, which will be expanded in Chapter 10, of the major categories of influence may now be stated: *Coercion* occurs if one party intentionally and successfully influences another by presenting a credible threat of unwanted and avoidable harm so severe that the person is unable to resist acting to avoid it. For a threat to be coercive, it is necessary (but not sufficient) either that both parties know that the person making the threat has the power and means to make it good or that the person threatened believes the other party to have the power, a belief of which the other party is aware. The threatened outcome of harm may be of many types: physical, psychological, economic, legal, and so on.

Manipulation is a catch-all category that includes any intentional and successful influence of a person by noncoercively altering the actual choices available to the person or by nonpersuasively altering the person's perceptions of those choices. The essence of manipulation is getting people to do what the manipulator intends by one of these two general means. Some scholars have regarded the absence of manipulation as essential for informed consent. Jeffrie Murphy, for example, holds that the reason for securing informed consent is to respect autonomy, "which means in this context making sure that he [the subject] is not merely *used* or *manipulated.*" He contends that "The information required for informed consent is that level of information which is necessary in order to eliminate or at least radically reduce the possibility of the subject's being *manipulated* by the physician."[31] Murphy discusses only informational manipulation, whereby the structure or perception of choices available to a patient is altered by managing information so that the person does what the physician intends. Although we would deny that informational manipulation exhausts the scope of manipulation, we do agree with Murphy that this form of manipulation is a major problem for informed consent.

Finally, *persuasion* is restricted to influence by appeal to reason. Persuasion is the intentional and successful attempt to induce a person, through appeals to reason, to freely accept—as his or her own—the beliefs, attitudes, values, intentions, or actions advocated by the per-

suader. Persuasion is always a nonclandestine form of interpersonal influence; the persuader openly puts forward reasons for accepting or adopting what is advocated. All choices made and acts performed on the basis of persuasion are noncontrolled, and, if the other conditions of autonomous action are satisfied, autonomous.

An appeal advanced in a storm of double-talking rhetoric that attempts to induce agreement or consent by making a person confused or disoriented does not qualify as an appeal to reason; it is an attempt to manipulate by altering perceptions through means other than reason. Similarly, manipulation, and not persuasion, occurs if an influence agent creates or otherwise brings about contingencies that subsequently come to function as reasons for another to act as he or she desires. We will expand on these themes in Chapter 10.

Although there is no need for informed consent requirements that would circumscribe persuasion per se, in many contexts in which informed consent is solicited, persuasion is difficult to distinguish from various forms of manipulation. (Similarly, it is often difficult—but also necessary—to distinguish certain more controlling forms of manipulation from coercion, which is always completely controlling.) For example, when information must be presented to patients and subjects, the presentation itself will contain, to varying degrees, nonsubstantive elements like tone, manner and order, word choice, time and setting, and the appearance, style, and character of the presentation agent. These need not be, but could be, potentially manipulative, as can the substantive content of an argument.

In general, it is not possible to specify with precision where persuasion ends and some types of manipulation begin. These concepts are too ill-formed; and ordinary language, philosophy, and the social sciences all fail to provide exact boundaries. Fortunately, we have no need to stipulate such precise boundaries. More important is to be able to distinguish substantially noncontrolling influences, as distinct from all other forms of influence. People generally choose and act in the face of competing wants, needs, familial interests, legal obligations, persuasive arguments, and the like. Many of these factors are influential without being to any substantial degree controlling. From the perspective of informed consent, we need only establish general criteria for the point at which substantial noncontrol is imperiled. This task too is reserved for Chapter 10.

Is Authenticity a Necessary Condition?

Thus far we have presented three conditions of autonomous action. But are they actually sufficient, and might not some other condition(s) be necessary?