Journal of Medicine and Philosophy, 37: 226–254, 2012
doi:10.1093/jmp/jhs016
Advance Access publication on May 2, 2012

Voluntary Consent: Why a Value-Neutral Concept Won't Work

ALAN WERTHEIMER*
National Institutes of Health, Bethesda, Maryland, USA

*Address correspondence to: Alan Wertheimer, PhD, Department of Bioethics, National Institutes of Health, Bethesda, MD, USA. E-mail: Wertheimera@cc.nih.gov

Some maintain that voluntariness is a value-neutral concept. On that view, someone acts involuntarily if subject to a controlling influence or has no acceptable alternatives. I argue that a value-neutral conception of voluntariness cannot explain when and why consent is invalid and that we need a moralized account of voluntariness. On that view, most concerns about the voluntariness of consent to participate in research are not well founded.

Keywords: *autonomy, coercion, consent, voluntariness, undue influence*

I. INTRODUCTION

It is uncontroversial that (barring exceptional circumstances) it is unethical to enlist people in research without their valid consent. It is also uncontroversial that valid consent to participate in research must be voluntary. As the Nuremberg Code puts it, "The voluntary consent of the human subject is absolutely essential" (Nuremberg Code, 1949). Or, in the words of the Belmont Report, "An agreement to participate in research constitutes a valid consent only if voluntarily given" (Belmont Report, 1979). Although Belmont maintains that "This element of informed consent requires conditions free of coercion and undue influence," it does not say, and it is not clear whether those conditions are thought to be sufficient to establish that a subject's consent is voluntary. And what constitutes coercion or undue influence? Does encouragement or persuasion by one's doctor compromise voluntariness? Does one consent voluntarily if one believes one has an obligation to do so? Or because one is paid? Or because participation in research is the only way to get needed medical care?

Here, we encounter considerable controversy. Many people advance or accept claims about the voluntariness of consent to participate in research

that are deeply puzzling and would seem to have little traction in most other consent contexts—commercial transactions, employment, sexual relations, litigation, or medical treatment. On some views, much consent to participate in research does not pass the test of voluntariness. By contrast, I will argue that many of the worries about voluntariness of consent to participate are not well founded and are rooted in a value-neutral account of voluntariness that cannot explain why and when consent is valid.

A few words about the *range* of issues that I will be discussing. As Appelbaum and colleagues have noted, most of the literature on consent to participate in research (along with consent to medical treatment) has focused on the cognitive dimension of valid (informed) consent—whether the relevant information is disclosed, how well it is understood, and whether the agent's reasoning is distorted by "internal determinants such as confusion, fear, or unreasonable hope" (Appelbaum, Lidz and Klitzman 2009, 30). On an expansive conception of voluntariness, all these factors might be thought to compromise the voluntariness of consent. Indeed, on some views, involuntariness refers to virtually any external or internal factor that compromises the validity of consent or the consenter's capacity to act as an autonomous agent. Although little turns on words, we will achieve greater analytical clarity if we distinguish between (1) internal cognitive or reasoning deficiencies and (2) external constraints that impact voluntariness. On this narrower conception of voluntariness, for example, the therapeutic misconception is a cognitive error that may compromise the validity of consent, but it does not render one's consent involuntary. Indeed, on this narrower conception of voluntariness, if A engages in deception or withholds important information, B's consent may not be valid, but A's action does not compromise the voluntariness of B's consent. In any case, our task here will be simplified if we assume for the sake of argument that subjects are fully competent, that they have been provided with all relevant information, that such information is well understood, that their reasoning is not distorted, and so on.

II. WORRIES ABOUT VOLUNTARINESS OF CONSENT

Before going further, it will be useful to identify the sorts of worries that have been expressed about voluntariness to participate in research. The first two worries were expressed to me personally.

Obligation

I was recently asked to be in the control group of a study of thrombosis. Participation involved a blood draw and a 45-min survey that involved some cognitive tests. When the investigator thanked me (too) profusely for participating, I remarked that I believed and had written that people have an obligation to participate in such trials (Schaefer, Wertheimer, and Emanuel,

2009). She replied, "I hope you didn't consent for that reason." Further conversation revealed that she thought that my consent would not have been sufficiently voluntary if I consented because I felt obligated to do so.

Persuasion

My physician encouraged me to participate in a Phase III randomized trial of the timing of chemotherapy for asymptomatic chronic lymphocytic leukemia. The trial (which never completed for lack of subject accrual) was designed to determine whether it was more or less efficacious to wait for symptoms to appear before starting chemotherapy. After signing the consent forms, I was asked whether I'd also be willing to participate in a "quality of life" study that was piggybacked onto the clinical trial. The consent administrator said, "I know this is coercive, but we'd really like you to do this." I don't think that she *really* thought it was *coercive*, but it was clear that she was worried that trying to persuade me to participate was bringing too much pressure to bear or that I would be concerned not to disappoint my physician and that my consent would not be sufficiently self-directed.

Difficult or Desperate Background Conditions

It is frequently argued that those in desperate conditions such as illness or extreme poverty may have no acceptable alternative but to participate in research given their need for payment or medical care and that the voluntariness of their consent is, therefore, suspect. This concern is particularly acute in much international research in developing countries.

Incentives

Many think that offering payment as an incentive to participate in research compromises or has the potential to compromise the voluntariness of a subject's consent. Roberto Abadie writes that the notion of a "paid volunteer" is an "oxymoron"—"How can someone simultaneously be paid to do something and to do it voluntarily"? (Abadie, 2010, 45). Etymology may confuse. We do sometimes use the word "volunteers" to refer to those such as hospital volunteers who work without pay, but we do not think that being paid to do something renders one's action *involuntary*. But even if payment does not entail involuntariness, many think that offers of payment can constitute coercion or undue influence if payment is sufficiently large and, thereby, compromises the voluntariness of consent (Macklin, 1989).

Perceived Threats

It is uncontroversial that if a doctor should threaten to abandon a patient if he does not agree to participate in research, then the patient's decision is coerced and involuntary. Robert Nelson and Jon Merz go further. They argue

that "The fear of loss of health care benefits or of retribution for refusal to participate render any given decision coerced, *regardless of the researcher's intention*," and this may be so even if the patients are vulnerable to "*imagined threats* that would not be credible or could be resisted under other circumstances or by other people" (Nelson and Merz 2002, v. 75).

III. VOLUNTARINESS AND VALIDITY

Bioethicists worry about voluntariness to participate in research because we accept the principle that consent is valid only if it is voluntary. Call this the *validity requires voluntariness principle*. By valid consent, I refer to consent that is morally transformative, that is, consent that renders it permissible for another person to do that which it would be impermissible to do without such consent, for example, to have sexual intercourse, to provide medical treatment, to extract a kidney for transplant, or to use one as a research subject.

If we accept the validity requires voluntariness principle, it follows that if B's consent is involuntary, then B's consent is not valid. Let us refer to the claim that B's consent is involuntary as an *involuntariness claim*. Before we determine whether we should accept an involuntariness claim, it appears that we must first determine its "truth conditions," that is, the factors that would render it true.

Roughly speaking, there are two views as to the truth conditions of an involuntariness claim. One view maintains that an involuntariness claim is a (reasonably) straightforward empirical or value-neutral proposition. As Nelson and colleagues put it, "voluntariness is not a value-laden concept" (Nelson et al., 2011, 7). On this view, an involuntariness claim depends on "synchronic nonmoral facts" that can be settled by examining the agent's options, beliefs, and capacities at the time of consenting. A second view maintains that the truth conditions of an involuntariness claim include moral judgments, that such a claim is fundamentally moralized. I shall argue that the first—empirical or value-neutral—view should be rejected because it cannot helpfully explain when and why consent is valid. To do that, we must adopt some version of the second—moralized—view.

Within the territory of empirical or value-neutral accounts of involuntariness, some think that an involuntariness claim refers to something about B's mental state at the time B consents. So it might be thought that to do something involuntarily is to do it reluctantly or unhappily or unwillingly. One advantage of this preanalytic view of involuntariness is that it has some linguistic and phenomenological support. The disadvantage of this view is that the spirit with which one consents turns out to be irrelevant to the validity of consent. For example, the fact that B is reluctant to sign a contract has no bearing on its validity.

There are, however, at least two more plausible candidates for a value-neutral conception of involuntariness. In their recent important article, Robert Nelson et al. argue that a voluntary action be understood "in terms of two necessary and jointly sufficient conditions: intentional action . . . and the absence of controlling influences" (Nelson et al., 2011, 6). Although it is not entirely clear when influences are "controlling" on their account, they allow that in addition to intentional acts designed to control another person, controlling influences can refer to internal psychological states or external circumstances—"A person may feel controlled by severe illness [or the] lack of a basic resource" (Nelson et al., 2011, 9). For present purposes, the crucial point is that Nelson and colleagues avowedly seek a value-neutral conception of voluntariness. In particular, they argue that whether it is morally legitimate for someone to exercise a controlling influence has nothing to do with whether B's consent is voluntary.

Another popular empirical conception of involuntariness asserts that one chooses involuntarily if one's choice is made because there are *no acceptable alternatives*. Within the framework of this view, there can be disputes over details. Is nonacceptability objective or subjective? Does a person act involuntarily if she *believes* or *feels* that there are no acceptable alternatives—even if she is mistaken? How bad must an alternative be for it to be the case that one has no *acceptable* alternative? Does one act involuntarily if one believes or if there are no *morally* acceptable alternatives? For present purposes, I put these interpretive issues aside. I will assume that we can identify when an agent is subject to controlling influence or has no acceptable alternatives and that this can be done (more or less) in value-neutral terms. The crucial question is whether a value-neutral conception of involuntariness can explain when one's consent is invalid. I will argue that it cannot.

Now I do not deny that involuntariness *can* be sensibly defined in value-neutral terms. I will argue, however, that although the controlling influence and no acceptable alternative views have considerable intuitive and scholarly support, neither (or any other comparable) value-neutral view can provide a plausible account of the sort of involuntariness that compromises the validity of consent. The claim that B's consent is involuntary *in a way that yields invalid consent* must include reference to the moral legitimacy of the actions of those who induce or solicit the agent's consent.

IV. THE LEGAL APPROACH

In their very helpful recent article, Paul Appelbaum and colleagues suggest that we turn to the law for guidance on voluntariness. They correctly note that for *legal* purposes, a decision "is presumed to be voluntary if no evidence exists that someone else has unduly influenced it or coerced the person deciding" (Appelbaum, Lidz, and Klitzman, 2009, 32). From this

perspective, a decision is not regarded as involuntary if it is driven by the agent's own values and preferences or the agent's *circumstance*, such as poverty, illness, or, in medical cases, the lack of "alternative treatment options." Indeed, the law regards B's decisions as voluntary even if A has exercised controlling influence over B or has made other alternatives unacceptable to B—if A's actions are legitimate.

Appelbaum and colleagues have pointed us in the right direction, but they do not provide an argument as to why the legal model is appropriate for *ethical* analysis of voluntariness. After all, there may be special reasons for the law to adopt a strict view as to what compromises the validity of consent. For example, the legal system may be concerned to reduce litigation or *ex post* challenges to agreements and so provides only narrow bases on which to claim that one's consent is involuntary and therefore invalid. Moreover, whatever the law says about involuntariness, there might be some reason to think that from a *moral* perspective, if one consents because one has no acceptable alternatives or is subject to a controlling influence, then one's consent is involuntary and one's consent is not morally valid. We need to determine whether the legal model of voluntariness is defensible from a moral point of view.

V. WORDS, CONCEPTS, AND MORAL FORCE

Words such as "voluntariness," "voluntary," "voluntarily," and "volunteer" have multiple legitimate meanings or usages. There is no reason to think that there is a univocal account of the proper use of the word. We do not need language police in this neighborhood so long as we understand what we are doing. Because we use voluntariness in different senses, it is important to see that its proper use in one sense does not entail its proper use in another sense. We *do* sometimes invoke the notion of voluntariness in a purely descriptive sense to refer to a sense of being unpressured and, in particular, unpressured by *others*. So B might say, "My decision to donate blood is completely voluntary; no one asked me to give blood." But it would not follow that B's consent is involuntary if someone had persuaded B to give blood by pointing to the current shortage. On the other hand, one might invoke voluntariness simply to deny that one was conscripted or forced to do something but not to deny that one had incentives for consenting. So one might say, "I wasn't drafted into the military; I volunteered for the money." And, as the advocates of a value-neutral account of voluntariness might argue, to say that an agent acts involuntarily may denote that the agent had no acceptable alternatives or that he has been subject to a controlling influence. Call this $involuntariness_{descriptive}$ (Burra 2010).

Although one can invoke voluntariness and its cognates simply to convey information about one's choice, situation, or motivation, the concept is often

invoked because it has *moral force.* In some cases, to say that an agent acted involuntarily implies that she should not be held responsible or liable for some of the consequences of her action. Call this *involuntariness* $_{responsibility}$. In addition, and of most interest to our present inquiry, to say that an agent acts involuntarily may imply or entail that her consent should not be treated as valid, that it does not give A permission or obligate B to perform some act. Call this *involuntariness*$_{consent}$. Now, it is logically possible that these three sorts of involuntariness claims always converge that involuntariness$_{descriptive}$ always entails involuntariness$_{responsibility}$ or involuntariness$_{consent}$. If this were so, then a straightforward value-neutral conception of voluntariness is all that we would need. But, as we will see, this is not so.

To see that involuntariness$_{descriptive}$ does not entail involuntariness$_{consent}$, let us widen the lens by considering a range of consent cases that do not involve consent to participate in research.

Nonmedical Cases

A "consent decree"—note the name—is a judicial order confirming agreement by a defendant to cease activities alleged by the government to be illegal in return for an end to the charges. The Environmental Protection Agency may threaten to bring criminal or civil charges against, say, British Petroleum (BP), unless the latter agrees to stop certain activities and pay a specified financial settlement. In return, BP will not be required to admit guilt or fault. Surely, BP would not agree to the settlement in the absence of a threat of suit or prosecution and the belief that the chances of losing big are sufficient to make it rational for it to settle.

Suppose we say that a party's consent is involuntary when she has no acceptable alternatives or is subject to a controlling influence. What are we going to say about the voluntariness and validity of BP's consent decree? We can say that BP has acceptable alternatives or that it is not subject to controlling influences and so its consent is voluntary$_{descriptive}$ on these value-neutral accounts of voluntariness, but then it would seem that consent to participate in research is rarely involuntary$_{descriptive}$. If we assume that BP's consent is involuntary on a value-neutral account, there appear to be three alternatives: (1) We can accept the validity requires voluntariness principle and say that BP's consent is involuntary and therefore invalid, (2) We can reject the validity requires voluntariness principle and say that BP's consent decree is involuntary but valid, and (3) We can accept the validity requires voluntariness principle and say that that although BP's consent is involuntary$_{descriptive,}$ it is voluntary$_{consent.}$ In my view, (1) reaches an implausible conclusion about the validity of the consent decree; (2) reaches a plausible conclusion, but it requires us to explain how consent could be both valid and involuntary; and (3) also reaches a plausible conclusion, but it poses a challenge: It requires

us to explain *why* and how BP's consent is *voluntary* given that B has no acceptable alternatives or is subject to a controlling influence.

Consider plea bargaining. A prosecutor may offer a defendant the choice between going to trial and risking a severe punishment and pleading guilty to a lesser charge and the certainty of a lesser punishment. Given that a defendant cannot be compelled to waive his right to a trial by jury, we must ask whether the structure of plea bargaining is compatible with voluntarily consenting to plead guilty. In plea bargaining, the prosecutor has intentionally created a choice situation such that a defendant may have no acceptable alternative but to plead guilty. Moreover, the prosecutor may exercise "controlling influence" over the defendant. One could claim that defendants have acceptable alternatives or that prosecutors do not exercise controlling influence, but then one would be hard pressed to distinguish the situation of defendants from cases where one thinks people do not have acceptable alternatives or are subject to controlling influences. So assuming that a defendant's guilty plea is involuntary on a value-neutral account, we have three alternatives: (1) We can accept the validity requires voluntariness principle and say that the defendant's plea should not be accepted, in which case the defendant may have to stand trial; (2) We can reject the validity requires voluntariness principle and say that the defendant's plea is involuntary but that we should accept the guilty plea nonetheless; and (3) We can accept the validity requires voluntariness principle and say that even though the defendant's plea is involuntary$_{\text{descriptive,}}$ it is voluntary$_{\text{consent}}$ because the prosecutor's proposal is not morally illegitimate or because the prosecutor is proposing a more lenient punishment than he has a right to pursue (Wertheimer, 1987, Chapter 7).

To further pursue this issue, consider the distinction between *extortion* and *hard bargaining*.

Extortion

Sam threatens to break the windows of John's restaurant unless John agrees to hire Sam's garbage removal company. John signs a contract.

Hard bargaining

Tom, who is John's long-standing supplier of beef, tells John that he must agree to a 50% price increase or find another supplier. There is no other available supplier of acceptable quality beef. Tom signs a contract to pay the increased price for 12 months.

What should we say about the voluntariness and bindingness of the two agreements? Many would regard John's agreement with extortionist Sam as involuntary and not morally binding but regard John's agreement with Tom as voluntary and binding—and this is so even though John had no

acceptable alternative but to sign *both* contracts and it is arguable that both Sam and Tom exercised a controlling influence over John's decision. With respect to John's agreement with Tom, we have three alternatives once again: (1) We could treat John's agreement as involuntary and therefore invalid; (2) We could treat it as valid despite its involuntariness$_{\text{descriptive}}$; and (3) We could say that John's agreement is voluntary$_{\text{consent}}$ and valid even though it is involuntary$_{\text{descriptive.}}$

The contrast between *extortion* and *hard bargaining* illustrates an important point at the center of disputes about voluntariness. Serena Olsaretti argues that the factors that make choices involuntary when carried out in response to a coercive threat are the very factors that make other types of limited choices involuntary, namely that "the agent makes the choice he makes because he has no acceptable alternative" (Olsaretti, 1998, 54). Consider the choice between accepting employment and starving.

> The alternative faced by the man who hands over the money when threatened with a gun is to be killed; the alternative of a worker who sells his labour power for whatever price is to remain unemployed and eventually starve. *The relevant condition which undermines voluntariness in the first case is also present in the second, namely, the absence of an acceptable alternative.* (Olsaretti 1998, 72, emphasis added)

Olsaretti presumes what has to be shown. Although it is true that the gunman's victim has no acceptable alternative to turning over his money, it does not follow—and I would deny—that *this* factor explains why his consent is involuntary and invalid. If I am right, then it also does not follow that the worker's consent to sell his labor power rather than starve is also involuntary and invalid.

Olsaretti might concede that whereas the gunman's victim has been *coerced* because he has been threatened with a harm, the worker has not been coerced because no one has threatened to make him worse off if he declines an employer's offer. Nonetheless, she would argue that the worker's decision is comparably *involuntary* because, in her view, it is the absence of acceptable alternatives and not the presence of coercion that renders one's decision involuntary. Once again, the issue is not whether we can reasonably *say* that the worker's decision is involuntary. We can. The issue is whether the fact that a worker has no acceptable alternatives (on some plausible view of that notion) but to accept employment entails that we should regard the worker's consent as involuntary in a way that would render her consent invalid. I will argue that it does not.

To see this, consider a standard employment contract. At Time-1, B, an unemployed lawyer who had experience painting houses as a teenager, signed a contract with A to paint A's house for $10,000 and A agreed to pay B a $5,000 deposit up front. B signed because she had no acceptable alternative; she was unemployed and had a child to support. At Time-2, and prior to the time at which B was to begin painting, C offers B a job with C's law

firm. Is B's consent at Time-1 morally valid? There is no dispute about the facts. We can agree that B had no acceptable alternative but to sign. If someone wants to claim that B's agreement was therefore involuntary that does not solve the practical problem. We must still determine what to say about the moral status of B's agreement with A. Here, I will simply assert that whatever we say about the voluntariness of B's decision we should *not* say that B's agreement with A is morally invalid and not binding. Although the law will not demand that B paint A's house, B is certainly under a legal and moral obligation to return the deposit and, perhaps, to compensate A for other losses.

Or consider *Dating*.

Dating

A and B have been dating for a while but have not had sexual relations. A tells B that unless she agrees to have sex, he will break off the relationship. Although B would prefer not to have sexual relations with A, other things being equal, she regards breaking off the relationship with A as an unacceptable alternative and so reluctantly consents to sex.

Now depending on further facts, we might think that A's behavior is caddish and manipulative. Still, even if we accepted B's claim that she had no acceptable alternatives, we would not think that B's consent is invalid. We would surely not think that A was guilty of rape or any other form of sexual crime, given that A is under no obligation to continue the relationship on B's preferred terms.

Taking Stock

The principal lesson of the previous examples is that there are numerous nonmedical contexts in which we do and should treat an agent's consent as valid even if the consent would be regarded as involuntary on a value-neutral account of voluntariness. As we have seen, this poses a trilemma: (1) We can accept that validity requires voluntariness and say that B's consent is invalid, (2) We can reject the validity requires voluntariness principle and say that B's consent is valid nonetheless, and (3) We can accept the validity requires voluntariness principle and adopt a moralized or nonvalue-free account of voluntariness and say that B's consent is voluntary.

Medical Cases

I now want to consider a set of medical cases in which it has been or may be claimed that consent is involuntary. Let's start with consent to medical treatment. Consider the case in which a patient consents to medical treatment, such as chemotherapy, breast surgery, dialysis, or amputation because the only alternative is death. Can such a patient give valid consent to treatment?

Arthur Caplan has argued that it is "hard to imagine" that those facing "certain death" can "be said to exercise informed consent . . . since the very fact of imminent death limits the realities of choice to doing anything that a physician offers as holding any hope" (Caplan 1997, 35). Now, we do distinguish between elective surgery and nonelective surgery, and I readily concede that there is a sense in which we regard the former as more voluntary than the latter. But assuming that what Caplan calls "certain death" refers to what will happen without treatment and that treatment offers a nontrivial prospect of survival, we have the same three options that we saw in the previous section: (1) We could accept the validity requires voluntariness principle and say that a patient's consent is involuntary and that her consent is therefore invalid, (2) We could reject the validity requires voluntariness principle and say that the consent is involuntary but that the consent is valid perhaps because something less than voluntary consent is sufficient to authorize necessary medical treatment, and (3) We can accept the validity requires voluntariness principle and say that the consent is voluntary and valid even though the patient has no acceptable alternative.

Given these options: (1) seems wholly unsatisfactory. If the patient's consent is invalid, we could decide (1a) that it is impermissible to provide life-saving treatment, which would be absurd or (1b) that it is permissible to treat *without the patient's consent*, in which case it would be pointless to seek her putatively invalid consent except, perhaps, as a ritual of politeness or to determine whether she regards death as an acceptable alternative to treatment. Both (2) and (3) allow us to treat the patient's consent as valid. But (2) exacts a high price as we would have to violate the principle that it is impermissible to treat a patient without her voluntary consent and it is unclear why this less than voluntary consent should render such treatment permissible. I believe that (3) best captures our practice as well as our linguistic and moral intuitions. We certainly act as if a patient's consent is necessary to rendering it permissible for the physician to provide such treatment even though patients are choosing treatment in order to avoid what would otherwise be certain death. But (3) presents its own challenge. It is not difficult to show that the patient's consent has not been *coerced* given that she has not been threatened with an adverse consequence by another person. But it is more difficult to show that her consent can reasonably be regarded as voluntary when she truly has no acceptable alternative.

Consider "voluntary euthanasia." Dutch euthanasia legislation states that euthanasia is only permissible if it is based on a voluntary request made in a situation of unbearable suffering to which there are no alternatives. Somewhat ironically, the legislation presupposes that a "voluntary request" is not only compatible with a situation in which there are "no alternatives"; it claims that such requests should be honored *only if* there are no alternatives. Martin van Hees suggests that the legislation contains an internal contradiction.

If the patient has no acceptable alternatives, then "a request for euthanasia cannot be said to be voluntary" (Van Hees, 2003, 62). Along similar lines, Neil Campbell writes that "If the pain and suffering are by definition unbearable, then it seems clear enough that the decision to die is not freely chosen but is compelled by the pain" (Campbell, 1999, 243). And, he argues, just as a prisoner who is tortured for information is not responsible for revealing state secrets when he is subject to "excruciating pain," the person who asks that his life be ended is not responsible for his decision. Campell and van Hees do not deny that a decision to request that one's life be ended under such conditions might be *rational*. But they do deny that it can be voluntary.

Once again, words do not matter much. Whatever words we use, what is often referred to as "voluntary euthanasia" would still stand in contrast to paradigmatically involuntary euthanasia to which the agent does not consent at all. If one claims that voluntary euthanasia is really involuntary, we would *still* have to decide whether there is an important moral distinction between "requested" and "nonrequested" euthanasia, and we would still have to decide whether and when such requests are sufficient to render euthanasia permissible. No view about the concept of voluntariness will resolve this substantive moral issue.

Setting aside linguistic issues, is there a substantive moral justification for not permitting voluntary euthanasia on grounds of its alleged involuntariness? (There may be other reasons not to allow voluntary euthanasia.) Van Hees thinks that euthanasia might be defensible on consequentialist grounds as a means by which to end unavoidable suffering but that it cannot be defended on deontological or autonomy-respecting grounds. On his view, since suffering itself vitiates the voluntariness of the patient's request, the suffering "undermines one's autonomy and thus also the moral legitimacy of the request" (Van Hees, 2003, 63). I disagree. As the themes of movies such as *Million Dollar Baby* and *Whose Life is it Anyway?* serve to illustrate, we always have to ask what it is to act autonomously within one's circumstances, horrible though they may be. We think that patients can autonomously refuse life-preserving treatment or request that such treatment be halted under dire circumstances and that their choices should be respected. Similarly, it may well be argued that patients can autonomously request euthanasia under truly awful conditions or, even further, that the ability to make such a request is a *fundamental exercise of one's autonomy*.

Consent to organ donation raises a different set of concerns about voluntariness. Some are concerned that an "opt-out" system in which society can take one's organs upon one's death unless one explicitly opts out is not compatible with a requirement of voluntary informed consent. Others have argued that paying people for giving up an organ may violate a requirement of voluntary consent if an impoverished person feels he has no acceptable alternative but to accept the payment.

Maryam Valpour makes a different point. She argues that a person who feels morally obligated to donate an organ does not do so voluntarily or autonomously. "If obligation is experienced in such a way that a donor feels he/she cannot refuse donation even though he/she does not want to donate, then that consent is coming close to substantially controlled . . . and, therefore, nonautonomous" (Valpour 2008, 198). This is puzzling. First, what does it mean to say that one feels one "cannot refuse donation even though [one] does not want to donate"? We need to distinguish between "simple wants" and "all things considered wants." Consider a woman who wants to have intercourse at a given time because she wants to become pregnant. Although she may not desire sex, qua sex, at that time, she may very much want to have intercourse—all things considered—when the things to be considered include her desire to become pregnant. Similarly, a person who donates because he feels an overwhelming obligation to do so *does* want to donate *all things considered* when the things to be considered include one's belief that one should donate and one's motivation to do what one thinks one should do. If one had been wrongfully manipulated into feeling that one had an obligation to donate, then one's decision might be regarded as "substantially controlled" and "nonautonomous." But if someone should decide to donate to his sibling because after all, it's "my brother," I see no reason not to regard his consent as voluntary and autonomous even if—perhaps especially if—he believes that he has no morally acceptable alternative.

Arthur Caplan shares Valpour's concern. He says that emotionally related donors "may find it impossible to give their consent freely . . . because they feel coerced . . . by the nature of the obligations that they see as defining their relationship to the person in need." On his view, "if consent is to be valid, then those giving it must feel free to say no" (Caplan 1997, 117). Otherwise the transfer of an organ is "taking, not giving, battery not altruism." Caplan seems to think that such decisions are driven by emotions that control or overwhelm one's decision making rather than being based on genuine moral beliefs and affection. But I see no reason to think that emotionally (and biologically) related donors who choose to donate out of love or a sense of obligation are not doing so voluntarily or that we should be suspicious about the validity of their consent.

A more interesting problem arises when a person wants to donate for self-interested reasons—broadly understood. Consider this case.

A needs a kidney. B is A's brother. B is a good match. A and B have never gotten along and, other things being equal, B would refuse. But both A and B do get along with other members of their family, and the other members bring considerable pressure to bear on B to donate. This pressure does not overwhelm B's capacity to think about things rationally and despite the family's urging, he's still not convinced that he has an obligation to donate. Nonetheless, B values his relationship with his family and is concerned that

this relationship would be damaged if he refused. B also fears that his parents might alter their will if he refused. B consents.

I see no reason not to regard B's consent as valid and voluntary. He is making a reasonable judgment about his familial and financial interests in a context in which no one has threatened to violate his rights (he has no *right* to inherit from his parents) if he refuses.

VI. THE MORAL FUNCTIONS AND TRUTH CONDITIONS OF INVOLUNTARINESS CLAIMS

In considering numerous nonmedical and medical examples, I have argued that we do not and should not treat B's consent as invalid just because B consents because he has no acceptable alternative or because he is subject to controlling influence. But even if these value-neutral accounts of involuntariness do not entail that one's consent is invalid, $\text{involuntariness}_{\text{descriptive}}$ may serve other moral functions. For example, it may excuse B from some ascriptions of responsibility or blame. Suppose that B accepted a job as a stripper in order to support herself and her children but that someone has criticized her for participating in an activity that objectifies and exploits women. B might claim that she had no acceptable alternative.

I want now to suggest that the truth conditions of an involuntariness claim vary with its moral functions. In other words, the same set of facts may justify us claiming that an agent's consent is $\text{voluntary}_{\text{consent}}$ but $\text{involuntary}_{\text{responsibility}}$. In particular, to say that B's consent is $\text{involuntary}_{\text{consent}}$ or renders B's consent invalid may require that the pressures are generated by another person and that they are illegitimate or coercive. Olsaretti suggests that B's consent to A's performing a medical procedure when she has no acceptable alternative may still entail that B bear the "integral" costs of the decision, including being unable to sue A for battery or invading her body without his consent (Olsaretti 2008, 118). With respect to *these* moral consequences of her decision, Olsaretti concedes that B's consent is or, perhaps, might as well be voluntary or at least valid. By contrast, Olsaretti suggests that the fact that B does not have acceptable alternatives may entitle her to complain about the fee if A should over-charge B even though B may have consented to pay that price.

The link between voluntariness and responsibility is actually a bit more complicated. The truth conditions for an involuntariness claim that limits ascriptions of liability may differ from the truth conditions of an involuntariness claim that limits ascriptions of praise. Suppose that C (a child) is drowning in a pond. A is a nonswimmer, but sees that B is about to pass by without making an effort to save C. If A coerces B to rescue C at the point of a gun, B deserves no praise for rescuing C. If, however, B rescues C because B believes he has no acceptable *moral* alternative, then he deserves whatever

praise would be consistent with the rescue being voluntary (if rescuing is obligatory, then great praise may not be in order.) At the same time, if B's spouse should criticize B for ruining his clothes in the course of rescuing C, B may claim "I had no choice." And so his rescuing was arguably voluntary with respect to praise but involuntary with respect to responsibility for costs.

By contrast, and as we have seen, there are many cases in which we do and want to treat an agent's consent as valid or morally transformative even though she has no acceptable alternatives or is subject to controlling influence (as in the BP and plea bargaining cases). An old advertisement for a motor oil filter had a mechanic saying, "You can pay me now [for the oil change] or pay me later [for repairing your engine]." With respect to the relationship between voluntariness and validity, we can invoke a consideration of moral factors before or after the assessment of voluntariness. First, we could adopt a value-neutral account of voluntariness and then abandon the validity requires voluntariness principle. In this case, the moral analysis as to whether the consent is valid will come *after* establishing that consent is involuntary. Second, if we want to retain the validity requires voluntariness principle, then we must adopt a moralized account of voluntariness$_{\text{consent}}$ under which the voluntariness of the agent's action turns on the morality of the actions of *others*. In this case, moral analysis would come *before* the determination as to whether the agent acts voluntarily. Either way, we cannot assess the validity of consent without appeal to moral analysis. A value-neutral account of voluntariness cannot explain why and when we should regard consent as valid.

Olsaretti comes close to accepting the possibility of a moralized approach to voluntariness when she writes that there is nothing mysterious "in claiming that different features of someone's . . . making a choice are justifiably relevant for different types of response" (Olsaretti 2008, 118). But although Olsaretti has argued that different constraints on choices may justify different moral upshots or responses, she has not gone so far as to argue that the *voluntariness* of those choices may vary with those upshots or responses. And that is the view that I wish to defend or at least render plausible.

The Deeper Theory

I have argued that the legal model of voluntariness recommended by Appelbaum and colleagues turns out to be a plausible model of morally valid consent. On the legal model, consent is regarded as involuntary and invalid only if the pressures on the consenter are morally illegitimate, although there may be factors orthogonal to voluntariness such as deception and incompetence that defeat the validity of consent. In particular, the law treats B's consent as voluntary and valid if A has made an offer of a benefit to B if B consents to X, or if B feels obligated to consent (or do) X, or B consents because B's circumstances are such that B feels she has no acceptable

alternative. Indeed, and although this is complicated, if a third party should wrongfully pressure B to consent to A's doing X, the law may treat B's consent as voluntary and valid so long as A, himself, has not put illegitimate pressure on B and is not taking unfair advantage of B's situation. So if B's wife should tell B that she will leave him unless he enrolls in A's drug rehabilitation program, B's consent will be treated as valid even if B's wife was wrong to so threaten him (of course, her threat may have been perfectly legitimate) and even if A knows that B is enrolling because of the threat.

The problem is this. Recall BP. Someone might say, "I accept that BP's consent decree is valid or enforceable, but I deny that BP's consent was *voluntary*." Even if it makes sense to treat consent as *valid* in a variety of circumstances in which an agent's choice is involuntary$_{\text{descriptive}}$ because she has no acceptable alternative or is subject to controlling influence, how can we treat such consent as *voluntary*? To show it is plausible to regard such consent as voluntary is a genuine challenge.

It will prove best to respond to this challenge by beginning with the reasons for treating consent as *valid*. There are two related arguments for treating B's consent as valid in the face of no acceptable alternative or controlling influence. The first argument is consequentialist; the second argument is rooted in respect for the agent's autonomy. From a consequentialist perspective, there are two reasons why we generally insist that certain sorts of transactions or interventions are permissible if but only if the agent consents. First, we regard the agent's voluntary consent as *necessary* to protect the person from interventions that are not wanted and that do not advance the party's interest. Second, and of paramount but often overlooked importance, we regard the agent's consent as *sufficient* (other things being equal) to authorize transactions or interventions or create binding obligations when doing so will advance his interests. From a consequentialist perspective, it would be a serious mistake to regard consent as nontransformative just because the consenter had no acceptable alternatives or is subject to controlling influence.

To see this, consider background situations such as illness or poverty in which the agent has no acceptable alternative but to consent to (or request) an intervention. We hope that we do not find ourselves in such circumstances, but, when we do, we want to be able to authorize interventions or create permissions that will advance our well being *within* that situation. We hope not to have cancer, but, if we do, we want to be able to authorize chemotherapy or surgery in order to avoid a premature death. We hope not to be unemployed and on the brink of starvation, but we would like to be able to give consent that will be recognized as valid to otherwise unattractive employment options (or even attractive options) if we were to find ourselves in such a position. We can make similar claims about consenting to otherwise unattractive options (e.g., pleading guilty) if one were subject to the controlling influence of a prosecutor. If we were to say that these factors compromise voluntariness and if we also insist on the validity requires

voluntariness principle, we would be preventing people from engaging in welfare-improving transactions.

But here we encounter a problem. If we should treat an agreement as valid whenever it is in a party's interest (at that time) to do so, then it may be argued that agreements made in response to extortionate threats should also be treated as valid.

A, a gunman, tells B that he will kill B unless B gives A $1,000. B does not have $1,000 but is willing to sign an IOU for $1,000.

Once B finds himself in this situation, it may well be in B's interests to be able to enter into a binding agreement with A. For if A realizes that the IOU would not be regarded as binding, A is more likely to kill B. So it would seem that consequentialist considerations tell in favor of treating B's IOU as valid and binding, in which case the consequentialist strategy seems to prove too much. But it does not prove too much. Although there might be occasional benefits to treating such agreements as valid and binding, a general policy of treating such agreements as valid would seriously threaten the stability of the basic framework of rights and liberties within which the possibility of consensual mutually advantageous transactions takes place. As a general matter, it is in one's ex ante interest to be the recipient of legitimate proposals, such as an offer to perform life-saving surgery, but contrary to one's ex ante interests to receive extortionate threats.

Now consider the deontological or autonomy-respecting argument for regarding many cases of consent as valid when the consenter has no acceptable alternative or is subject to controlling influence. Although autonomy is a notoriously difficult concept, for present purposes, I shall simply stipulate that to be autonomous is to be in control of one's life and exercise self-determination. There is both a negative and a positive dimension to respect for autonomy. On the one hand, we require that people's consent is voluntary and informed in order to protect their negative autonomy from interventions to which they do not genuinely agree. On the other hand, we respect a person's positive autonomy when we allow him to authorize interventions or bind himself to do something. And so, we fail to respect a person's positive autonomy when—with excessive concern for his negative autonomy—we do not allow him to authorize interventions or facilitate binding agreements. Because it is impossible to simultaneously maximize respect for both dimensions of autonomy, it is difficult to get the balance right. Consider a patient with a painful and terminal illness who is considering voluntary euthanasia. To emphasize her positive autonomy by allowing her to authorize such euthanasia is to risk allowing her to make such a decision when she is less than fully competent. On the other hand, to emphasize her negative autonomy by not allowing her to authorize euthanasia because she is less than fully competent may be to condemn her to continued suffering. The present point is that any plausible conception of autonomy must be sensitive to both dimensions of autonomy.

When we focus on the positive dimension of autonomy, it becomes clear that any plausible conception of self-determination operates within a certain conception of the world in which people find themselves—a world that contains poverty, prosecutions, civil suits, illness, and, in the worst case, unbearable and unavoidable suffering. In addition, the world in which we exercise our own self-determination is also defined by the rights of others who are exercising their autonomy. Recall John, who has received an extortionate threat from Sam and a demand for a higher price from beef supplier Tom. John can reasonably want *ex ante* protection from such extortion (by a state that punishes such acts) and *ex post* protection by a policy of not treating his consent as valid. But as much as John might prefer that Tom continue to sell him beef at the lower price, Tom has a right to sell only at a higher price and John cannot reasonably claim that his right to operate his restaurant prohibits Tom from making such demands. Any account as to when John's consent is morally transformative must be rights sensitive in this way. Similarly, in Dating, although B would prefer to continue her relationship with A without sexual intercourse, A has the right not to continue dating on B's preferred terms. That is the context in which she exercises her own autonomy and must decide which alternative is preferable.

We can reach a similar conclusion by adopting a contractualist theory with respect to the validity of consent. Although such a theory can be modeled in different ways, we can ask what principles about the validity of consent would be adopted in something like a Rawlsian original position in which people are choosing principles for a nonideal world, not knowing what positions they would occupy in that world. The contractors would understand that illness, accident, and even unjust background conditions may place them in situations in which they can improve their situation only if they have the ability to authorize someone else to do something to them or for them. They would *not* adopt a conception of valid transactions that renders one's consent invalid or nontransformative whenever one is in a position in which one has no acceptable alternative or is subject to controlling influence. And they would certainly not adopt a conception of valid consent that did not permit them to act on their considered moral views about their obligations. And so both consequentialist and deontological arguments support the view that we should sometimes treat consent as valid even when the consenter is subject to controlling influence or has no acceptable alternative.

VII. FROM VALIDITY TO VOLUNTARINESS

If we should regard many cases of consent as valid even when given under very difficult conditions or when the consenter is subject to controlling influence, that still leaves open the question as to why we should regard such consent as *voluntary*. To (re)use an example, can we defend a conception of voluntariness that allows us to say that John's response to extortionist Sam

is involuntary, whereas his response to beef supplier Tom is voluntary? Appelbaum and colleagues note that the law assumes such a view but that does not show that this view is conceptually or morally coherent.

There are two types of responses to this challenge. What we might call an incompatibilist response accepts a value-neutral view of involuntariness and simply gives up the attempt to square voluntariness with validity and abandons the validity requires voluntariness principle. On this view, we would still need principles for distinguishing the involuntary consent that is valid from the involuntary consent that is invalid. We might accept the claim that voluntary euthanasia is actually involuntary, but then argue that we can treat a person's involuntary consent to euthanasia as valid. And the same might be said about BP's consent decree, guilty pleas, some employment contracts, and the choice of chemotherapy. This is a coherent view. Moreover, if we have to choose between rejecting the validity requires voluntariness principle and accepting morally unacceptable conclusions (e.g., that one cannot authorize life-saving medical treatment), it is more important that we reach the right conclusions than that we retain our allegiance to the validity requires voluntariness principle. Nonetheless, the marriage between voluntariness and validity is one to which we are linguistically and intuitively strongly committed, and so we should accept a divorce here with great reluctance. We can't easily say, for example, that a defendant's guilty plea is involuntary but valid, given the constitutional guarantee of a right to a trial by jury. We can't say that BP's agreement with the government is valid but made under duress, given that a contract made under duress is voidable. If, as a last resort, we have to abandon the validity requires voluntariness principle, so be it, but we should first if we can defend a conception of voluntariness that allows us to say that some cases of involuntariness$_{\text{descriptive}}$ can plausibly be regarded as voluntary$_{\text{consent.}}$

There are at least three compatibilist responses to this challenge. The first compatibilist strategy appeals to the phenomenology of voluntariness. It matters to us whether our decisions reflect our will *within a situation* as contrasted with cases in which our decisions are driven by the will of *another person* who is seeking to determine our decisions *in ways that we would reject.* The previous clause is important. Suppose that A persuades B that B should do X or offers B an incentive so that B will consent to do X. Although A is getting B to do what A wants B to do in both cases, B need not resent the fact that he is choosing what A wants B to choose, so long as B does not regard the means by which A is attempting to influence B's choices as illegitimate or contrary to what B would prefer. Our experience of choice is sensitive to whether others are acting within their rights. If C tells B that she will leave B unless B agrees to participate in A's drug addiction rehabilitation program, B might feel angry at C, but he is unlikely to resent A's willingness to allow him to enroll. Of equal importance, there are many situations in which B feels as if a choice is *his* choice, not because he has

numerous alternatives from which to choose but principally because the choice is *not someone else's*. Something like this may be true in voluntary euthanasia. Although I believe that this phenomenological strategy lends some support toward reconciling voluntariness$_{\text{consent}}$ with involuntariness$_{\text{descriptive,}}$ I don't believe that it is strong enough. Among other things, there will be cases in which people fail to make the distinction just noted. For example, although A may be acting within his rights in attempting to exercise controlling influence over B's decision (as in plea bargaining), B may still feel that his choice is forced or involuntary, even if he also wants his agreement to be treated as valid.

A second compatibilist strategy appeals to the well-known hierarchical or two-tiered view of the will, first made famous by Harry Frankfurt. Although there is a sense in which we always do what we prefer, there is another (allegedly more important) sense in which acts are autonomous or voluntary when they are compatible with our more reflective (higher or underlying) preferences, whereas involuntary acts are not. On this view, "autonomy is conceived of as a second-order capacity of persons to reflect critically upon their first-order preferences, desires, wishes, and so forth and the capacity to accept or attempt to change these in light of higher order preferences and values" (Frankfurt 1971, 11). On Frankfurt's view, a drug addict's decision to consume a drug is not autonomous or voluntary if she has a higher order preference not to consume drugs, even if it is informed, rational (under the circumstances), and uncoerced. Along similar lines, we might argue that consent that appears *not* to be autonomous or voluntary because the agent has no acceptable alternative or is subject to controlling influence at the time of the decision can be considered as autonomous and voluntary if the agent has a higher order or reflective view about the terms on which her consent *should* be regarded as valid, for example when one's consent is driven by circumstantial pressures or is *not* driven by illegitimate pressures.

A third and related compatibilist view argues that the *values* that underlie our concern with the voluntariness of consent would *not* support a value-free conception of voluntariness that is combined with the validity requires voluntariness principle. After all, we want valid consent to be voluntary because many typical cases of involuntary consent are not likely to promote our interests or respect our autonomy. But as I have shown, if we combine a value-free conception of voluntariness with the validity requires voluntariness principle, we will reach conclusions about the validity of consent that defeat the very values that underlie our concern with voluntariness in the first place. We will be unable to give valid consent in cases where doing so advances our interest or respects our autonomy.

Taking Stock Again

In addition to the three compatibilist arguments just discussed, there may be other strategies for defending the compatibilist position. Note that these

arguments are consistent with the claim that consent is involuntary$_{\text{consent}}$ *only if* one's consent is also involuntary$_{\text{descriptive}}$ according to the best value-neutral account of involuntariness. But the compatibilist arguments also maintain that one's consent is voluntary$_{\text{consent}}$ when—as a rough approximation—there is good moral reason to regard one's consent as valid. In other words, involuntariness$_{\text{descriptive}}$ is necessary but not sufficient condition for involuntariness$_{\text{consent.}}$ If we opt for a single value-neutral account of voluntariness, then the only viable alternative is to abandon the validity requires voluntariness principle. That would be a pyrrhic victory for the advocates of the value-neutral conception of voluntariness because the involuntariness of consent would then be of reduced moral significance. Voluntariness would have limited bearing on the validity of consent. By contrast, if we adopt a moralized account of the truth conditions of voluntariness, then we can have our cake and eat it, too. We can accept the validity requires voluntariness principle and yet reach plausible conclusions about the conditions under which consent is valid. Although I have not given a knock down argument for a moralized account of the coherence of the concept of voluntariness$_{\text{consent,}}$ let us assume that the cumulative weight of the arguments discussed (or some additional arguments) renders such a view plausible. We can then pose this question: When should we regard consent to participate in research as involuntary?

VIII. VOLUNTARINESS OF PARTICIPATION IN RESEARCH

Setting aside cases in which people are literally coerced or conscripted into research, my analysis of the concept of voluntariness suggests that many of the worries about the voluntariness of consent to participate in research are misplaced. In general, we should assume that the principles as to what constitutes involuntary consent in other areas of life also apply to consent to participate in research. If circumstances such as poverty do not compromise validity of consent to employment and serious illness does not compromise validity of consent to accept medical treatment, there is no reason to think that such circumstances—by themselves—should compromise the validity of consent to participate in research. There may be good reason to worry about some sorts of pressures on prospective participants that do not apply in other contexts, but, in general, we have good consequentialist and autonomy-respecting reasons not to regard consent to participate in research as involuntary or invalid just in case one has no acceptable alternatives or is subject to controlling influence.

Let's start with Caplan's claim that severe illness itself is coercive and compromises the validity of consent to participate in research. Caplan does not regard this remark as mere semantic hyperbole. University of Pennsylvania researchers were engaged in a Phase I trial of gene transfer therapy for ornithine transcarbamylase deficiency, a rare metabolic disorder. Jesse Gelsinger,

18 years old, had a mild form of this disease that was controlled with a low-protein diet and medication (Steinbrook 2008). A more severe form of the disease was virtually always fatal with infants. The question arose as to whether the trial should be conducted with adults who could consent to participate but to whom the research posed serious risks or with infants with the more serious form of the disease who could not consent (their parents could consent for them) but who would otherwise die soon anyway and to whom, in that sense, participation posed little risk of being worse off. Caplan, the university's resident bioethicist, argued that the experiment should not be conducted with infants not just because the infants could not consent but because the parents of dying infants are incapable of giving informed consent: "They are coerced by the disease of their child" (Stolberg, 1999). Jesse Gelsinger consented to participate and died.

Needless to say, I am not arguing that Caplan's flawed view of coercion and voluntariness was the sole or even a major factor in the decision to use Jesse Gelsinger as a subject or that our evaluation of the decision should be made *ex post* by reference to the actual events. The case is much too complicated for any such judgment. I am arguing that adopting the view that illness itself is coercive and that such illness driven consent is involuntary and invalid is not without consequence. And this is particularly so in cases such as some Phase I oncology research in which a patient is prepared to consent to participate in research because she sees participation in research as offering the best option for successful treatment (although we may reasonably worry that some participants may overestimate the probability that they will benefit from participation).

Caplan aside, I suspect that few actually adopt the view that serious illness precludes voluntary and valid consent to participate in research when participation offers the best treatment option. The problem is that institutional review boards are more likely to adopt this flawed no acceptable alternative view of involuntariness when nonmedical conditions such as poverty lead prospective subjects to think that participation in research is their best alternative. Nelson and colleagues argue that constraining situations such as poverty can lead to deprivations of voluntariness that are "morally problematic," although they also say that it is "problematic" to deny people with the capacity for voluntariness the opportunity to participate on grounds of their circumstances (Nelson et al., 2011, 9). At the end of the day, describing consent as "problematic" won't help. We have to decide whether we should treat such consent as valid. For reasons I have given, I believe that we should regard consent as voluntary and valid if—as I have assumed from the outset—subjects can make a competent and rational evaluation of the risks and benefits of participation.

Second, and related to the previous point, let us consider the claim that offering incentives to participate in research—particularly large incentives—can compromise the voluntariness of participation via coercion or undue

influence. When Ruth Macklin says that the question of how large a payment constitutes a coercive offer is one for "which no clear answer is forthcoming," she implies that sufficiently large offers can be coercive (Macklin 1989, 3). I have argued elsewhere that offers of payment—however large—cannot be coercive because coercion ordinarily requires a threat of harm (Wertheimer and Miller, 2008). Eschewing the language of coercion, Eleanor Singer and Mick Couper argue that monetary incentives should be considered "unduly influential" if "they induce participants to undertake risks, they would not be willing to accept without the incentive" (Singer and Couper, 2008, 50). As a general proposition about voluntariness and incentives, this simply can't be right. There are many risks or inconveniences or burdens that people will not accept without monetary incentives, to wit, almost every form of employment and, in particular, risky forms of employment, such as lobster fishing and logging. On the assumption that people are competent and informed decision makers, unless we can find some principle that justifies us saying that offers of payment to participate in research compromise the voluntariness of consent but that offers of payment to engage in other activities do not do so, then it's hard to see why the offer of payment compromises the voluntariness of consent to participate in research. I have no idea as to what such a principle would look like.

In their recent article, Nelson and colleagues argue that offers of payment are not morally problematic if the risks are low but become problematic as "(1) Risks are increased to an elevated level, (2) More attractive inducements are introduced, and (3) The subjects' economic disadvantage or lack of available alternatives . . . is increased" (Nelson et al., 2011, 9). Given that we regard people's consent to assume (1) elevated risks in the employment sector as valid in the face of factors such as (2) and (3), in the absence of argument as to why we should think that the acceptance of risks in research participation should be regarded as different, I see no reason to think that offers of payment to participate in research are problematic in a way that would render them involuntary or invalidate consent.

Now, I readily grant that offers of payment would compromise the *validity* of consent if such offers *distort* a subject's ability to evaluate the benefits and risks of participation. This, I believe, is the proper realm of undue influence. On a mistaken but popular "sliding scale" view, some understand "undue influence" as a weaker cousin of coercion. But undue influence is better understood not as referring to constraints on voluntariness (the realm of coercion) but, rather, to the impact of offers on a subject's decision-making capacity. There is undue influence if subjects are likely to overweight the benefits of participation or underestimate the risks. But there is no undue influence when subjects are capable of making rational evaluations of benefits and risks and come to the conclusion that participation in research is their best option.

Appelbaum and colleagues argue that the legal model of voluntariness implies that B's consent is invalid if A's proposal is illegitimate or immoral.

This is not quite right. A's proposal may be illegitimate for numerous reasons, but it would ordinarily not compromise the validity of B's consent unless A proposes to violate B's rights if B does not consent. Suppose that A offers B, a professional killer, $25,000 to kill C. We would not claim that B consents to the agreement involuntarily and is not responsible for his act of killing C even though B would not have killed C but for the incentive provided by A. Why? Because A is not proposing to violate B's rights (by not paying B) if B does not consent. We can make a similar point about bribery. If A offers B, a policeman, $100 if B will "tear up the ticket," we would not say that B acts involuntarily in accepting the bribe or that B is not responsible for doing so. So even if paying research subjects is a form of bribery—and I think it is not—it would not follow that those who accept the bribe are doing so involuntarily.

If the offer of payment as an incentive to consent to participate in research does not compromise the voluntariness or validity of consent, much the same can be said for the offer of medical benefits, such as novel interventions or free medical treatment or medical screening that would otherwise be unavailable. From the self-interested subject's perspective, the incentive of receiving needed medical treatment is, in principle, no different from the incentive of payment. In both cases, people choose to participate because they believe that it is in their all things considered interest to do so.

Although few people participate principally because they believe that they have an obligation to do so, for reasons discussed above, there is no reason to regard altruistic motivations as compromising the voluntariness of participation (Stunkel and Grady, 2011). It would be strange to say "I can't accept your consent to donate blood as voluntary and valid because you have told me that you are consenting because you feel morally obligated to do so." So, too, for participation in research.

Finally, let us consider three methods by which physicians may attempt to motivate patients to participate in research. It is uncontroversial that a physician may not threaten to stop treating or abandon a patient unless the patient agrees to participate in research. Interestingly, it is less clear why this is so. A physician has a right to decide who he's going to treat. Given this, why do we not regard a physician's proposal to stop treatment unless the patient consents to participate in research as akin to Tom's proposal to continue to supply beef only if John agrees to a higher price? It must be because there are reasons to circumscribe the reasons for which a physician may propose to stop treatment that do not apply in other commercial relations. For example, physicians may have a nonwaivable fiduciary obligation to their patients that does not hold between sellers and buyers. I will not spell out or defend this view of the physician–patient relationship in detail. Suffice it to say that if a physician did make such a proposal, a patient who consented under such circumstances would be doing so involuntarily not solely or even primarily because he had no acceptable alternative or is subject to a

controlling influence. The pressure on the patient may not be all that great. Rather, we would regard the consent as involuntary and invalid because we regard the physician's proposal to make continued treatment conditional on participation as illegitimate.

Interestingly, even if it is wrong for a physician to abandon a patient that he has begun treating, it does not follow that it would be wrong for a physician to make the start of treatment contingent upon participation in a trial. And so, it may be argued that a patient who consents to participation because the doctor has made the commencement of treatment contingent on such consent has done so voluntarily and that the consent is valid. Is this argument correct? On the one hand, physicians have considerable discretion with respect to the persons whom they will treat. For example, they can refuse to take on any new patients or patients who cannot pay for their services or who propose to pay via Medicaid. On the other hand, this discretion is not unlimited. We do bar physicians from excluding patients on certain grounds, such as race and national origin. Thus, although a physician can decline to enter into such a relationship on some grounds, he is not free to decline on any grounds at all. Once again, we have to make a judgment as to the legitimacy of reasons for which a physician can refuse to enter into a treatment relationship. Should we say that physicians should be barred from excluding patients because they do not wish to participate in clinical trials?

I'm not sure. At the National Institutes of Health, patients with rare conditions may get standard treatment by world-class experts in connection with some elements of research separate from the treatment (say, blood draws required for research but not required for treatment). They are not eligible to get the treatment without research participation. Here, a decision to make availability of treatment conditional upon research participation may be perfectly appropriate. If so, then patients who consent to such participation are doing so voluntarily. The present point is not to resolve the ethical status of such programs. The point is that the question as to whether consent to treatment conditional upon related participation in research is voluntary and valid will not be resolved by a conceptual analysis of voluntariness. It will be resolved by substantive moral analysis of the conditions under which it is permissible for physicians to make treatment conditional upon participation in research. For example, we may ask if the terms of the treatment are fair and whether they are consistent with an acceptable view of the physician/patient relationship.

Finally, let us consider persuasion. I noted above my own experience with being persuaded to join a study. Although this mild encouragement hardly compromised the voluntariness or validity of my consent, there may often be reason to worry about the effects of a physician's attempt to persuade a patient to participate in a clinical trial. I have argued that the criteria for voluntariness to consent to participate in research are not sui *generis*, and this is particularly so when subjects are healthy volunteers who are not recruited

by their treating physicians. But participation in research may be a sphere in which we may want to be sure that subject/patients receive special protection from verbal encouragement to reduce the probability that they are motivated by their fear of abandonment by their physicians or, less dramatically, because they fear disapproval by their physicians.

I think the problem is fear and not coercion, per se. Robert Nelson and Jon Merz argue that "The fear of loss of health care benefits or of retribution for refusal to participate render any given decision coerced, *regardless of the researcher's intention*," and this may be so even if the patients are vulnerable to "*imagined threats* that would not be credible or could be resisted under other circumstances or by other people" (Nelson and Merz, 2002, v. 75). This is puzzling. If the doctor has made no threats and if the patient's fears are unreasonable, then even though the patient *feels* coerced, it does not follow that he *is* coerced or is acting involuntarily in a way that would render his consent invalid. Nonetheless, even if a physician has not caused such fears, if he is or should be aware that the patient may have such fears, then even though the physician has not coerced the patient, he should not exploit those fears. Moreover, given that these fears may be reasonably widespread, there may be reason to give some prophylactic protection to patients from this particular form of pressure so that patients can be assured that physicians regard their interests as paramount and so that they—and the public—have less reason to worry that the vulnerability of patients is not being exploited.

Even if patients do not fear abandonment, is there reason to protect patients from consenting to participate in research because they seek their physician's approval? As a general principle, there is no reason to think that consent motivated by desire for approval is involuntary or invalid. After all, we do *many* things precisely because we seek the approval of others or fear their disapproval. This is the stuff of life. If B pledges to contribute to her neighbor's "Walk for AIDS" because B seeks her neighbor's approval, I would assume that he does so voluntarily.

But this is a place in which consent to participate in research might be different. It is important that patients be able to trust their physicians. The needs generated by illness and the inherently asymmetric character of the relation between patients and physicians is such that it might be desirable to protect patients against the exploitation of their desire for approval even if this means that their physicians cannot make perfectly benign efforts to persuade them to participate in clinical research. We might think that it is perfectly fine for physicians to recommend their preferred treatment option even if patients are likely to choose it because they do not want to risk disapproval. But here the physician's recommendation is or should be constrained by his fiduciary obligation to the patient. By contrast, we may be more reluctant to allow physicians to recommend participation in research if the fiduciary obligation is less applicable to the research context.

I said that it *might* be desirable to protect patients from the exploitation of their desire for approval. There will surely be some cases in which a physician attempts to persuade a patient to participate in research because he reasonably believes that participation is in the patient's medical interests or because he thinks that the patient would be prepared to make an altruistic contribution to medical science if he understood the importance of the contribution that he would be making relative to the risks and burdens of participation. It would be unfortunate if a prophylactic barrier against persuasion were to preclude all such attempts to solicit consent from patients. What should be done? Two points. First, I am inclined to think that the right approach here will be dictated, in large part, by empirical considerations. We need to know whether a policy of restricting attempts to persuade patients to participate in research would, on average, best serve the interests of their patients (I set aside the effect of research on the interests of future patients). No policy will get it right in every case, and we need data as to the frequency and quantity of foregone gains if doctors do not attempt to persuade and the frequency and severity of the harms to patients if doctors are permitted to persuade. Second, and as I have argued at several points, we will not settle such questions by reference to a value-free conception of voluntariness.

IX. CONCLUSIONS

The principal theoretical aim of this paper is not to resolve the question as to whether consent is voluntary and valid in one case or another but to argue that no value-neutral theory of voluntariness is going to answer such a question. If we adopt a value-neutral account of voluntariness along the lines of "no acceptable alternative" or "controlling influence," then we will be forced to abandon the validity requires voluntariness principle and will have to determine whether and when we should regard involuntary consent as valid. We can retain our allegiance to the validity requires voluntariness principle by adopting a moralized account of voluntariness in which the voluntariness of a subject's consent turns on the legitimacy of the means by which her consent is solicited. Either way, the question as to whether we should regard consent as valid will turn on moral analysis.

It may be argued, however, that my moralized account of involuntariness involves a form of circularity. On the standard picture, an involuntariness claim supports a moral claim about the validity of B's consent. On a moralized account, we effectively use a moral judgment as to whether B's consent should be regarded as valid to determine we should accept an involuntariness claim. And so, it seems that a moralized account of voluntariness involves a form of circularity: (1) If consent is involuntary, then consent is invalid; (2) If consent should be valid, then consent is not involuntary. The argument is not strictly circular because involuntariness requires involuntariness$_{\text{descriptive}}$ is a

necessary but not sufficient condition of involuntariness$_{consent}$. Still, the circularity objection is partly correct. But I don't think that this circularity can be avoided. At the end of the day, the determination as to whether B's consent should be regarded as invalid is a moral question to be resolved by moral argument and cannot be resolved by appeal to a value-neutral account of voluntariness.

ACKNOWLEDGMENTS

I wish to thank several colleagues at National Institutes of Health for helpful comments: Frank Miller, Joseph Millum, Seema Shah, and David Wendler. I also thank several members of the Department of Philosophy at the University of Vermont for comments and discussion: Terrence Cuneo, Tyler Doggett, and Randall Harp. The views expressed in this paper are those of the author. They do not necessarily represent any position or policy of the National Institutes of Health or the Department of Health and Human Services.

REFERENCES

Abadie, R. 2010. *The professional guinea pig.* Durham, NC: Duke University Press.

Appelbaum, P., C. Lidz, and R. Klitzman. 2009. Voluntariness of Consent to Research. *Hastings Center Report* 39, 30–9.

Belmont Report. 1979 [On-line]. Available: http://ohsr.od.nih.gov/guidelines/belmont.html (Accessed April 24, 2012).

Burra, A. 2010. *Coercion, deception, consent: Essays in moral explanation.* Dissertation, University of California-Los Angeles.

Campbell, N. 1999. A problem for the idea of voluntary euthanasia. *Journal of Medical Ethics* 25:242–4.

Caplan, A. L. 1997. *Am I my brother's keeper?* Bloomington: Indiana University Press.

Frankfurt, H. 1971. Freedom of the will and the concept of a person. *Journal of Philosophy* 68:5–20.

Macklin, R. 1989. The paradoxical case of payment as benefit to research subjects. *IRB* 11:1–3.

Nelson, R., and J. Merz 2002. Voluntariness of consent for research. *Medical Care* 40:9. Supplement, v69–80.

Nelson, R., T. L. Beauchamp, V. Miller, W. Reynolds, R. Ittenbach, and M. F. Luce 2011. The concept of voluntary consent. *American Journal of Bioethics* 11:6–16.

Nuremberg Code. 1949 [On-line]. Available: http://ohsr.od.nih.gov/guidelines/nuremberg.html (Accessed April 24, 2012).

Olsaretti, S. 1998. Freedom, force and choice: Against the rights-based definition of voluntariness. *Journal of Political Philosophy* 6:53–78.

———. 2008. Debate: The concept of voluntariness—A reply. *Journal of Political Philosophy* 16:112–21.

Schaefer, O., A. Wertheimer, and E. Emanuel 2009. The obligation to participate in biomedical research. *Journal of American Medical Association* 302:67–72.

Singer, E., and M. Couper. 2008. Do incentives exert undue influence on survey participation? Experimental evidence. *Journal of Empirical Research on Human Research Ethics* 3:49–56.

Steinbrook, R. 2008. "The Gelsinger Case." In *The Oxford Textbook of Clinical Research Ethics*, edited by E. Emanuel, Christine Grady, Robert Crouch, Reider Lie, Franklin Miller, and David Wendler, 110–20. New York: Oxford University Press.

Stolberg, S. G. 1999, Nov. 28. The bioetech death of Jesse Gelsinger. *New York Times* 6:137.

Stunkel, L., and C. Grady. 2011. More than the money: A review of the literature examining healthy volunteer motivations. *Contemporary Clinical Trials* 32:342–52.

Valpour, M. 2008. The live organ donor's consent: Is it informed and voluntary? *Transplantation Review* 22:196–9.

Van Hees, M. 2003. Voluntariness, suffering, and euthanasia. *Philosophical Explorations* 16:50–64.

Wertheimer, A. 1987. *Coercion*. Princeton, NJ: Princeton University Press.

Wertheimer, A., and F. Miller. 2008. Payment for research participation: A coercive offer? *Journal of Medical Ethics* 34:389–92.