Deception in Medical and Behavioral Research: Is It Ever Acceptable?*

DAVE WENDLER

National Institutes of Health

AMID ALL THE DEBATE SURROUNDING RESEARCH with human subjects, there remains a point of widespread agreement: researchers must respect subject autonomy. In addition, everyone agrees that respecting subject autonomy requires researchers to inform subjects honestly about the true nature of their research. The uninitiated can be forgiven, then, for drawing what seems like the obvious conclusion: everyone agrees that researchers should honestly inform subjects about the true nature of their research. Unfortunately, honestly informing subjects can present an obstacle to research and precludes certain kinds of research altogether. For this reason, and in spite of the widespread unanimity regarding the importance of subject autonomy, the question remains, Is subject deception in research ever ethically acceptable? In the present article I will attempt to answer this question by offering specific conditions on when subject deception in research is acceptable.

The tension between subject autonomy and subject deception arises most clearly when subjects are deceived, not for their own good, but for

*The ideas and opinions expressed in this article are those of the author and do not represent any position, policy, or procedure of the National Institutes of Health.

The Milbank Quarterly, Vol. 74, No. 1, 1996
Published by Blackwell Publishers,
238 Main Street, Cambridge, MA 02142, USA, and 108 Cowley Road,
Oxford OX4 1JF, UK.

the good of science or medicine; I will focus my attention on this specific issue (call it "no expected benefit" deception). The question of when it might be acceptable to deceive subjects for their own benefit is important, but importantly different, and I will not consider it here. Instead, I attempt to answer the narrower question of under what conditions, if any, is no expected benefit deception acceptable. To simplify things, I shall make two assumptions:

1. Subject deception is prima facie unacceptable because it violates the subject's autonomy (i.e., the capacity to determine one's own course of action).
2. Deception *can* be justified (i.e., deception is not, in principle, always wrong).

With these assumptions in place, I begin by considering the Ethical Principles of Psychologists and Code of Conduct (American Psychological Association 1992) (henceforth called "the principles"), which offer a notable exception to the relative lack of scrutiny that has been accorded the topic of specific conditions under which subject deception might be acceptable.

The Psychological Principles

The principles offer three conditions for acceptable deception (American Psychological Association 1992, §6.15; for similar guidelines, see Office for the Protection from Research Risks 1993). The first reads as follows:

1. Psychologists do not conduct a study involving deception unless they have determined that the use of deceptive techniques is justified by the study's prospective scientific, educational, or applied value and that equally effective alternative procedures that do not use deception are not feasible.

The second clause of this first condition requires that, in order for a given instance of subject deception to be acceptable, there must not be any equally effective and feasible nondeceptive alternatives to that research. The requirement of *equal* effectiveness raises a minor problem. If, as we are assuming, deception is prima facie unacceptable, then it

follows that we might prefer a nondeceptive study, even if it is somewhat *less* effective, because the harm incurred by the decrease in effectiveness might be outweighed by the good that is achieved in avoiding the deception. For this reason, let's drop the word "equally" from the first condition. This change raises the obvious question of *how much* of a decrease in effectiveness we should be willing to accept in order to avoid subject deception. To aid our decision, first consider the importance of subject autonomy and its relation to deception. Roughly speaking, the question to be answered here is, Why is subject autonomy important and, given that importance, how can violations of subject autonomy be justified?

For our present purpose, we can assume that individuals understand their own preferences better than anyone else does. To this extent, autonomy is important because it allows subjects to do what they prefer to do. Or, taking the same point from another perspective, the more others control the subject's actions, by deception say, the greater the chance that the subject will end up failing to do what she prefers to do all things considered (i.e., what she would choose if she had not been deceived). The principles make the following stipulation in order to avoid this possibility:

> 2. Psychologists never deceive research participants about significant aspects that would affect their willingness to participate, such as physical risks, discomfort, or unpleasant emotional experiences.

Generally speaking, no expected benefit *deception* takes place within the context of no expected benefit *research* (i.e., deception that is not intended to benefit a particular subject typically occurs within the context of research that is not intended to benefit that subject). The assumption behind condition 2, then, is that subjects decide whether or not to enroll in no expected benefit research based on the level of risk involved. Of course, subjects do not consider only the level of risk. They also consider, to varying degrees, the source of the study's funding, the goals of the study, the kind of people running the study, and so on. (Subjects often base their enrollment decisions on monetary compensation as well. However, such compensation is intended to "zero out" the inconveniences that participating in the study entails so that subjects can make their enrollment decisions based, as condition 2 assumes, on the risks involved.) These additional considerations highlight a crucial difference between deceptive and nondeceptive studies. Typically, one does

not regard the purpose of a study, say, as posing any risk to subjects. Potential subjects are informed of the purpose of the study and can decline to participate if they find that goal objectionable. Unfortunately, deceiving subjects about the purpose of the study eliminates this opportunity, thus raising the possibility that subjects will contribute to the achievement of a goal that conflicts with their beliefs, perhaps even deeply held moral or religious beliefs. Now although this would be of great concern to subjects, it is not something that the principles take into account: condition 2 addresses only *physical* risks and harms.

Briefly, we can think of harms as states of affairs that "contradict" the preference(s) of a subject. For instance, most subjects prefer to avoid pain and retain physical integrity. Therefore, states of affairs that involve pain, or the destruction of physical integrity, cause harm to the subject in question. The "contradiction" of moral preferences, in contrast, is standardly understood, not as a harm, but as a "wrong." For our purposes, we need not get enmeshed in the subtleties of this distinction. For, leaving aside the specifics of wrongs and harms, many people care about their moral and religious beliefs as much as they care about their physical well-being. Put roughly (if we assume a subjective account of preferences for the moment), subjects can be as upset by the contradiction of their moral preferences (not to mention their psychological preferences) as by the contradiction of their physical preferences. Prima facie, then, we have no reason to restrict our account of acceptable deception to *physical* harms, and thus I will understand "risks" broadly to include any possibility that a subject's preferences will be contradicted and that a subject will be upset as a result of his or her participation in research.

The second condition also includes the qualification to "significant" risks. However, without further argument, this qualifier is either redundant (if "significant" refers only to those aspects of the study that are relevant to subjects' willingness to participate) or unjustified (if "significant" refers to some subset of those aspects), and I will omit it as well (more on this below). Thus amended, the second condition ensures that, as long as the use of deception does not conceal any risks (understood broadly), it will not alter subjects' enrollment decisions (call this the "neutral risk condition"). In other words, as long as the neutral risk condition is met, we can assume that subjects who consent to a deceptive study would have done so even if that study had not been deceptive. As a result, the neutral risk condition allows us to conduct deceptive studies

while respecting subject autonomy in the sense of maximizing the chances for subjects to choose their preferred course of action (but see below). Therefore, if maximizing the chances for subjects to choose their preferred option exhausted the importance of autonomy, this neutral risk condition would end the need for any further discussion of acceptable deception. However, by requiring that deception be justified—even when the neutral risk condition has been met—the principles implicitly acknowledge the need for additional conditions because deception harms subjects even when it does not affect the choices they make. The reason for this, briefly, is that deception is prima facie immoral, not only because it decreases the chances that subjects will do what they prefer, but also because it involves one person controlling what another person does (without that person's consent). For this reason, any policy on acceptable deception must minimize the extent of deception even when the neutral risk condition has been met. The principles use three conditions to accomplish this end.

The second half of condition 1 helps minimize the extent of deception by limiting the number of *times* that subjects can be deceived: subjects may be deceived only when such deception is necessary for the study in question. Condition 2 further minimizes the extent of deception by limiting its *scope:* researchers may not deceive subjects about any *risks* presented by the study. Finally, the principles restrict the *duration* of deception by including a final condition on acceptable deception—call it the "mandatory debriefing" condition:

> 3. Any deception that is an integral feature of the design and conduct of an experiment must be explained to participants as early as is feasible, preferably at the conclusion of their participation, but no later than at the conclusion of the research.

With these conditions in place, the principles ensure that any use of deception in research is minimized to the greatest extent possible. This leaves us with the question of how we might justify the deception that remains.

The principles justify any remaining deception in the first clause of condition 1: "the use of deceptive techniques is justified by the study's prospective scientific, educational, or applied value. . . ." Because we are considering "no expected benefit" deception, this value will not benefit the subjects who are being deceived. Rather, we are considering cases

where subjects are being deceived for the (possible) benefit of others. In other words, the principles offer a straightforward consequentialist justification for violating subjects' autonomy by deception: the harm to subjects is justified by the potential (scientific and/or medical) benefit to others. This brings us to the first substantive problem with the principles. Consequentialist justifications are viewed with great skepticism. Perhaps nowhere is this skepticism more widely held than within the context of human subjects research, where appeals to the potential good of others has an unfortunate track record. Given this history alone, we would do well to consider approaches to acceptable deception that do not depend upon consequentialist justifications. In addition, as we shall see in the next section, justifying deception using a consequentialist approach requires a cost/benefit analysis of the harm that deception does. In the following section, we shall see that, at a practical level, such analyses are almost impossible to carry out. Taken together, the next two sections strongly suggest that we need to consider alternative approaches to acceptable deception.

The Cherek Studies

Investigators who hope to study human aggression face several potential ethical problems including, most notably, the question, How does one study aggression without having subjects harm one another? The most obvious solution is to *deceive* subjects into believing that they are harming one another. To take a fairly benign example of this approach, Don Cherek asks his subjects to try to accumulate points by typing into a computer (Cherek 1990). Cherek tells his subjects that their computers are connected to one belonging to another person, who may occasionally steal some of their accumulated points. In response, Cherek's subjects are given three options:

1. retaliate by stealing some of their "opponents'" points
2. briefly protect their own account against further thefts
3. independently continue to gain points

In spite of what Cherek tells his subjects, there is no other person and no other computer. Instead, the subject's computer is programmed occa-

sionally and arbitrarily to subtract points from his or her account. (Cherek then measures aggression as a function of how often subjects choose option 1 in response to these thefts.)

Determining the acceptability of this deception requires, the principles tell us, that we first determine whether there is an alternative way of doing the same research that is nondeceptive, feasible, and effective. Presumably there are nondeceptive alternatives. Cherek could, for instance, have a member of the research team replace the computer program by sitting at a second computer that is connected to the one belonging to the subject and typing in the responses, thus producing a nondeceptive study. Is this study feasible and effective? Take effectiveness first. Condition 1 tells us that, although the nondeceptive alternative need not be *equally* effective, the tradeoff in effectiveness cannot be too great. Is it? Presumably, the person entering the data will make more mistakes than the computer. For this reason, the alternative study will require more subjects, and even then may yield less powerful results. In addition, the alternative study will be somewhat less feasible: a second computer will be needed, and lines will have to be run from it to the subject's computer. Given this decrease in feasibility and effectiveness, and the fact that the deception involved in Cherek's original study is both relatively minor and seemingly presents no risks to subjects, one could easily conclude that the tradeoffs required by the alternative study are too great. And, if this is right, Cherek's deceptive study, according to the principles, is acceptable. In response, one could, perhaps just as easily, argue that Cherek's study offers the perfect example of a deceptive study with a feasible, and effective, nondeceptive alternative. Granted, the alternative introduces an extra variable, but the alternative study is, nonetheless, relatively straightforward, and the mistakes made by the computer operator could be isolated and controlled for.

Notice that both sides to this debate agree on the (potential) decreases in feasibility and effectiveness resulting from the nondeceptive study. Their disagreement centers on whether we should accept this cost in order to avoid the use of deception. Now according to the principles—and this is my point here—which side is right depends upon how much the deception "costs." In other words, in order to determine whether or not avoiding this use of deception justifies the corresponding decrease in feasibility and effectiveness, we must know how much the deception "costs." To see the difficulty of making this determination, consider a second deceptive study.

Alcohol Studies

Gathering accurate data on alcohol abuse is complicated by the fact that alcoholics often provide misleading information when asked direct questions about their alcohol use. To circumvent this problem, Fleming studied the acceptability of concealing the intent of alcohol abuse questionnaires by including questions about smoking, weight, exercise, and drug use. He explains: "The additional questions are included to HIDE the true purposes of the questionnaire" (Fleming 1989, 313). Fleming administered the "general health" questionnaire without first informing his subjects that the investigator's research interests were limited to the alcohol use questions (prior to this, he had conducted a preliminary "surrogate" study in which he informed subjects of the use of deception) (see Fost 1975). At the end of the study, Fleming debriefed his subjects and, in order to estimate the cost of the deception, asked them two questions:

1. Were they upset by the deception?
2. If so, would they still be willing to participate in such research again?

The results: one-third of the subjects were upset by the deception, but, of these, two-thirds supported the study and said they would be willing to participate again. Fleming cites the relatively high number of subjects who were willing to participate again, combined with the importance of accurate information on alcohol abuse, to show that these kinds of deceptive studies are ethically acceptable.

Fleming bases his argument for the acceptability of deception in this case on what looks like the best method for estimating its harm: deceive subjects and then ask them how much it bothered them. In general, subjects can answer this question in either of two ways. First, they could provide a *quantitative* assessment of the harm by deciding, roughly, How *much* did the deception bother me? Alternatively, subjects could be asked to assess directly the harm of the deception when balanced against the value of the study. For instance, they could be asked, Was the harm that you experienced by being deceived justified by the value of the study? The problem with this latter, comparative, approach is that very few subjects will know the value of the study. Hence, very few subjects will be in a position to *compare* the harm of the deception to the value

of the study. This brings us to the first option: have subjects make a *quantitative* assessment of the harm of the deception. Unfortunately, there are several reasons to think that subjects will not be able to do this.

Most important, people dislike viewing themselves as victims; and that, in essence, is what subjects are being asked to do. The degree to which you were harmed when your physician deceived you reduces directly to how badly were you victimized. Thus, it would not be at all surprising to find that, in the hope of avoiding victimization (or the recognition thereof), subjects fail to answer this question accurately. For this reason, the quantitative approach will systematically underestimate the harm that deception does. And, because this approach appears to provide the best estimate of the harm of deception, there may be no control study that could be carried out to gauge the extent of this underestimate. Therefore, this approach will require that we harm subjects without a clear sense of how much harm we are doing. Furthermore, in addition to being unable to estimate the harm of the deception, we still face the problem of comparing whatever estimates of harm we get with the estimated value of the study. In particular we have to ask, Who makes this comparison? The researchers? The possible beneficiaries? The subset of deceived subjects who were particularly bothered by the deception?

By definition, no expected benefit deception harms one group but (potentially) benefits a second group. Therefore, even if we could come up with an accurate estimate of both the harm of the deception and the value of the study, we would be left with no way of carrying out the required comparison. Finally, because these harm estimates are being made post hoc, each study will have to be performed at least once in order to estimate its harmfulness. Put differently, the principles' approach to acceptable deception requires that we first perform each study in order to determine whether or not performing that study is ethically acceptable. In sum, what looks to be the best method for carrying out the cost/benefit analysis required by the consequentialist approach is both harmful to subjects and practically impossible to carry out. These problems alone, not to mention the suspect nature of consequentialist justifications in general, provide good reason to consider alternative approaches to acceptable deception. In the next section, I begin to develop such an alternative by considering, in some detail, the exact nature of the harms of deception as well as several ways to avoid those harms.

Understanding and Avoiding the Harms of Deception

As we have seen, Fleming argues for the acceptability of his study by citing the percentage of subjects who were willing to participate again even after being informed that they were deceived. However, these same results tell us that one-ninth of Fleming's subjects were upset by the deception to the point of refusing future participation. This way of viewing his results is instructive. Fleming's subjects are told the truth about the nature of the study, including its risks (but see below), the lack of expected benefit, the procedures involved, and the possible alternatives. His deception involves the ostensibly minor point that the researchers were interested in a specific aspect of the subjects' health, rather than the subjects' health in general. And yet a significant percentage of the subjects were bothered enough to be unwilling to participate in such research again. Presumably, many of these subjects were upset, not by the fact that the researchers were actually interested in alcohol use, but by the use of deception itself. This conclusion supports a piece of common sense: at least some people are upset, and therefore harmed, by the mere fact of being deceived, independent of the nature of that deception. It follows that no matter how carefully we minimize the *extent* of subject deception, along the lines of the principles, say, justifying deception on consequentialist grounds nevertheless entails our harming subjects without their consent. I suspect that many will accept this conclusion, but will then support the continued use of deception by arguing that, because the deception is minimized by the principles' three conditions, the remaining harm to subjects will be minimal as well. This is an important mistake.

Subjects like those in Fleming's study who are upset by the use of deception per se are, presumably, subjects who place a high value on being in control of their lives. Therefore, taking control of these individuals' lives by deceiving them without their consent involves our contradicting one of their most important personal values. The existence of this, possibly serious, harm reveals that the principles are internally inconsistent. To see why, recall that the mandatory debriefing condition requires researchers to inform subjects of the use of deception by the end of the study. Unfortunately, this information will upset many subjects, thus contradicting the original second condition stipulation that acceptable deception not conceal any unpleasant emotional experiences. Similarly,

the use of deception contradicts the amended second condition stipulation that the use of deception not conceal any risks of harm (including the psychological harm of learning that you have been used by researchers). In sum, the fact that at least some subjects are upset by deception per se establishes that, even if the deception involved in a particular study does not conceal any risks that are present independent of that deception, the deception itself introduces a new risk into the study. Further, allowing studies that fail to inform subjects of this risk contradicts the primary tenet of research ethics, according to which subjects may not be put at risk without their consent (unless, perhaps, the research has the expectation of direct benefit to the subject).

The obvious way of avoiding this harm would be to treat the risk presented by the use of deception in exactly the same way that we treat all the other risks involved in research participation: inform subjects of its presence. This is precisely the position that I argue for below. However, before we come to that, I want to consider an alternative response to the present line of argument. One could view the present conclusion that subjects are harmed by deception, even when the principles' three conditions are met, as evidence that we need to minimize further the extent of allowable deception. An obvious way of doing this would be to stipulate that researchers may not deceive subjects about the most sensitive aspects of each study. In Fleming's study, for instance, subjects were deceived about the study's purpose. However, subjects may have strong views about what goals they are willing to contribute to (e.g., some subjects may not want to contribute to any work that looks at genetic differences between races or examines additional uses for fetal tissue). With this in mind, one might argue that Fleming's results teach us, not that we should inform subjects of the use of deception, but that we should never deceive subjects about the more sensitive aspects of a study.

Recall that the original wording of the second condition stated that researchers may not deceive subjects about "significant" aspects of a study. One way to understand the present line of argument is in terms of offering support for this qualification, with the suggestion that the purpose of a study should be understood as one of its significant aspects. Unfortunately, this way of understanding the present argument also reveals why it fails. First, not all subjects will agree on what aspects of a particular study are significant or of special moral concern. For instance, in reviewing Cherek's study, an institutional review board (IRB) presumably would judge the fact that subjects are competing against a com-

puter, rather than against another subject, to be insignificant from the perspective of informed consent. According to the present suggestion, then, deceiving subjects about that aspect of the study would be ethically acceptable. The problem with this is that some subjects may find playing computer games, for religious reasons perhaps, to be morally repugnant. Therefore, by allowing this particular kind of deception, we will be allowing these subjects to be seriously harmed without their consent.

In response, one could point out that these subjects are harmed because they have idiosyncratic beliefs and that we should treat such cases no differently than the possibility that subjects will have idiosyncratic *physical* reactions to some apparently innocuous aspect of a study. Some subjects, for instance, may have a violent reaction to the color of the walls in the treatment room. Typically, we accept such possibilities without requiring that researchers inform subjects of absolutely every aspect of a study that could conceivably present a risk to them, such as the color of the walls in the treatment room, the fabric of the investigator's sweater, the nurse's religious beliefs, and so on. To the contrary, the accepted view of informed consent is that subjects must be informed of only those aspects of the study that *reasonably* can be thought to pose a risk. Understood in this light, one might think that I am holding deceptive studies to a higher standard than nondeceptive studies—I am arguing that deceptive studies are unacceptable, in part, because of the possibility of idiosyncratic responses by subjects.

In general, researchers should assess which aspects of their studies can reasonably be thought to pose a risk to subjects, based on the most specific information available concerning the subjects who will enroll in their studies. Roughly speaking, this process occurs at three levels. At the first, and most specific, level, risks should be assessed based on any available information concerning the *particular individuals* who will enroll in the study (e.g., information gained from past contact with subjects who plan to enroll, their charts, and so forth). Investigators who do not have such individualized knowledge should make their risk assessments at the second level, by considering what aspects of the study can reasonably be thought to pose a risk to the *class of subjects* that will be participating in the study. For instance, researchers studying Alzheimer's disease may not have personal knowledge of the individuals who will enroll in their studies. They should know, however, that individuals with Alzheimer's disease face unique harms as a result of their condition, and

they should inform the subjects of these risks. When nothing is known about either the particular individuals or the class of subjects who will enroll (e.g., studies involving "normal" volunteers), researchers should evaluate risks to subjects at the third level, in terms of what the *reasonable person* would regard as a harm.

In essence, the argument that we are considering makes the point that there is no fourth level at which subjects are informed of all aspects of a study (e.g., the color of the walls of the treatment room) that could possibly pose a risk. Leaving aside the practical difficulties, informing subjects of every detail that could pose a risk to them would serve to confuse, rather than to inform. There are simply too many *possible* sources of harm. Given that there is no fourth level of risk assessment, one might concede that deceiving subjects about "insignificant" aspects of a study presents a risk to subjects who have idiosyncratic beliefs concerning those aspects. However, this possibility is a consequence of the nature of informed consent in general. Therefore, it no more offers a reason to require that subjects be informed of the use of deception than the fact that some individuals may have a violent reaction to yellow walls provides a reason to inform subjects of one's decorating schemes. Now, although this line of argument has a certain plausibility—and, indeed, I suspect that many accept it as justifying the risks of deception—it misses a crucial point.

We can grant this much: we do not require that subjects be informed of all the possible aspects of a study that could pose a risk. Instead, in the way just outlined, we require that subjects be informed of all aspects of the study that could reasonably be thought to present a risk to them. But, notice that, even though there is no fourth level of risk assessment, the informed consent process does not stop after all the risks that are assessed from the first three levels have been enumerated. In addition, subjects are given the opportunity to reveal any idiosyncratic concerns (e.g., allergies) that might cause some aspect of the study to be risky to them in particular. Now, when it comes to *non*deceptive studies, it makes sense to expect subjects to be aware of any idiosyncratic concerns (e.g., specific allergies, a morbid fear of yellow walls) and to reveal those concerns to the investigator during the screening process. Unfortunately, the same process cannot be applied to deceptive studies. It would, for instance, be unfair to expect subjects who regard computer games as immoral on religious grounds to reveal this concern during the screening of Cherek's study for the simple reason that Cherek's subjects are explicitly

told that they are competing against another subject (not against the computer). The problem here is that deceptive studies effectively eliminate the possibility that subjects will reveal idiosyncratic concerns during the screening process regarding the concealed aspect of the study; the deception provides them with an explicit reason to conclude that those concerns are not relevant to the study at hand.

Furthermore, arguing against the requirement of informing subjects of the use of deception by focusing on the *nature* of that deception is somewhat beside the point. As considered earlier, some subjects are harmed by deception per se, independently of the nature of that deception. For this reason, not informing subjects of the use of deception places at risk not only those subjects who have idiosyncratic concerns relating to the concealed aspects of the study, but also those subjects who are harmed by deception per se. I conclude that deception, without consent, presents serious risks to subjects even when that deception is maximally minimized. Finally, although we have been focusing on the ways in which deception presents a risk to subjects who are deceived, we should not lose sight of the fact that the harms of deception do not stop there. Once subjects enrolled in nondeceptive studies learn that researchers systematically deceive *some* subjects, it will be difficult to convince them that they are not being deceived. Investigators may argue all they want that these subjects are not being deceived. Unfortunately, subjects who are aware that deception is sometimes used in research without the consent of subjects will have no reason to believe such statements. This is a very disturbing consequence of deception: ultimately it might affect studies (e.g., oncology and AIDS studies) where subject trust is literally a matter of life and death.

In addition, although I won't consider it in detail here, a similar point can be made concerning the corrosive effects of deception on the trust that is invested in science and medicine in general. Medicine and science are highly specialized fields, which depend, in large part, upon the trust of the public and other researchers: no one has the time, or the expertise, to check the results of all other researchers. However, once the public and other researchers learn that some researchers are sometimes deceptive, they will have no way of delimiting the scope of this deception, and thus we will have no way of containing the costs of that deception. In sum, the harms of subject deception in research are more widespread than is often acknowledged and may threaten the very foundations of research. Fortunately, there is good news. The present argu-

ment does not imply that deception must be prohibited altogether. Rather, it suggests that subjects must be informed of the use of deception. Briefly, adopting this approach universally would block the undermining of trust: subjects (and others) would know when they face a deceptive study, and the public (and other researchers) would know when scientists and physicians are using deception. Furthermore, this approach would allow potential subjects who are especially bothered by deception to avoid deceptive studies altogether and allow all others to consent to being deceived.

Second-Order Consent

The most straightforward way of apprising subjects of the fact that a particular study is deceptive is to require precisely that information as part of the informed consent process. Of course, requiring a statement that informs subjects about the exact nature of the deception would effectively eliminate certain kinds of research. Therefore, let us first consider the feasibility of apprising subjects of the use of deception, without informing them of the nature of that deception (call this second-order consent, or SOC). Adding this requirement to the (amended) principles yields the following account of acceptable deception:

1. The use of deception must be justified by the study's prospective scientific, educational, or applied value, and any effective nondeceptive alternatives must not be feasible.
2. Subjects must never be deceived about aspects of a study, like risks, discomfort, or unpleasant emotional experiences, that would affect their willingness to participate.
3. The deception must be explained to subjects as early as is feasible, preferably at the conclusion of their participation, but no later than the conclusion of the research.
4. Subjects must be informed of the use of deception prior to their enrollment in the study.

The statement of SOC must inform subjects of the use of deception and should explicitly offer the opportunity for debriefing. Here is a statement that meets these two conditions:

> SOC: You should be aware that, in order to complete this study, several of its details cannot be accurately described to you. For this reason, these details have been intentionally misdescribed. However, the investigator will be happy to accurately explain these aspects of the study to you at the end of your participation (end of the study). In addition, you are free to choose not to participate if you do not like this use of deception, or for any other reason, and your refusal will not be held against you in any way.

The use of the word "details" to describe the misdescribed aspects of the study is intended to reflect the restrictions that the (amended) second condition places on any acceptable deception. Recall: this condition states that subjects should never be deceived about aspects of a study "that would affect their willingness to participate, such as risks, discomfort, or unpleasant emotional experiences." For the purposes of informed consent at least, those aspects of the study that would not affect a subject's decision to enroll can be described as "details." As an alternative, we could explicitly signal the satisfaction of the second condition by saying, for example, "A safety committee has reviewed this study and has determined that this use of deception does not pose any risks to you." I have not included such a statement because it seems simply to raise the specter of nonexistent risks and thus may result in needless refusals to participate. This approach is consistent with the generally accepted view on informed consent that requires subjects to be informed of the risks involved in a particular study; subjects need not be informed of the *lack* of additional risks. SOC covers deception that results from researchers making false statements to subjects, that is, as the result of researchers misdescribing certain aspects of a study. Misdescription is not the only way that a study can be deceptive, however; it can also mislead when researchers *withhold* information from subjects.

We can think of deception as communication (understood broadly to include written, spoken, and behavioral communication) or, more to the present point, the lack of communication which reasonably can be expected to result in some subjects developing a false belief (or beliefs) about a study they are in, or are going to enroll in. To see how SOC should handle deception that results when information is withheld from subjects, consider first the standard example of doing so, namely, placebo trials. A typical (blinded) placebo trial randomly assigns some subjects to receive a drug, and others to receive placebo, while withholding from subjects the information about which they are receiving. Further-

more, active steps are taken (e.g., placing drug and placebo in identical capsules, double blinding) to ensure that subjects do not discover the truth. Clearly, then, placebo trials *withhold* important information from subjects, and subjects often end up with a false belief about what they are taking. Nevertheless, such withholding of information is widely accepted because placebo trials explicitly inform subjects that this is being done. For this reason, placebo trials are not deceptive; the reasonable expectation is that subjects will *not* develop false beliefs *as the result* of researchers withholding this information. Instead, subjects will remain uncertain about what they are receiving. (The false beliefs that subjects often develop about what they are taking are not the result of deception; they are the result of subjects' guesswork.)

Notice that placebo trials would be deceptive if they did not include an explicit statement that information is being withheld. For, given subjects' background assumptions that physicians prescribe active drugs, it would be reasonable to expect subjects to develop the belief that they are receiving a drug. Put in general terms, withholding information is deceptive when, without information to the contrary, subjects' (not necessarily reasonable) background assumptions could reasonably be expected to lead them to false beliefs. Of course, subjects' background assumptions may not be known to the investigator or the IRB. In that event, IRBs and researchers will have to make their best assessment of what would prove deceptive on a case-by-case basis. Previously I pointed out that deception effectively eliminates the possibility that subjects will reveal idiosyncratic concerns that are relevant to the aspect of the study being concealed. For this reason, unclear cases should be treated as (possibly) deceptive and thus should include a statement of SOC. In summary, the lesson of placebo trials is twofold:

1. Given subjects' background assumptions, withholding information can be as deceptive as positively misdescribing aspects of a study.
2. Studies that are deceptive as the result of withholding information are acceptable as long as they include an explicit statement along the lines of SOC.

Now to clarify how the withholding version of SOC is applied in practice, consider a specific example.

Imagine that an investigator wants to see whether or not giving drug X to males increases their level of aggression. As is often the case in such

studies, telling potential subjects of the goal of this study would almost surely invalidate its results. What options does the investigator have? First, the researcher could deceive subjects. For instance, the investigator could tell subjects that the purpose of the study is to test whether or not drug *X* has any effects on memory. In that case, the study would be clearly deceptive and hence would have to include a statement of SOC (as well as consider any risks that might arise, directly or indirectly, from this deception). This way of handling deception should be clear. Now consider the alternative possibility of simply withholding this information from subjects. The most obvious way of doing so would be to describe the study using an overly general statement to the effect that the researchers are looking at the links between drug *X* and behavior in general. This is a common strategy for studies of this kind. Is it deceptive? As we have seen, the test is this: Is it reasonable to expect that some subjects will develop a false belief concerning the nature of the study as a result of this withholding of information? In most cases, the answer here will be yes. Most subjects who are told that researchers are looking at the link between drug *X* and behavior in general will conclude, falsely, that the researchers are not focusing on a *specific* form of behavior. One solution, then, would be to replace the overly general statement with a specific statement that omits the particular behavior being studied, saying, for example, "The researchers are investigating the link between drug *X* and a particular form of behavior." In certain cases, at least, this statement would not be deceptive and, hence, would not require an explicit statement that information is being withheld. Alternatively, researchers could retain both the overly general statement and the explicit statement that information is being withheld.

One could argue, in response, that overly general statements should not require SOC because this withholding of information is not as harmful as explicit misdescription. In the case of overly general statements, the true purpose of the study "falls under" the general statement (e.g., aggressiveness is one form of behavior). Therefore, subjects know that what, in fact, is being studied is a possible candidate for study, and they have the opportunity to reveal idiosyncratic concerns about that aspect (e.g., concerns about displaying aggressive tendencies). The overly general statement does avoid this harm and thus will be *less* harmful than other kinds of deception. Unfortunately, this is only one of the harms that deception causes. Previously we saw that deception also harms both subjects who are bothered by deception per se and subjects who are in-

volved in nondeceptive studies, and that it damages as well the general trust in science and medicine. Given these other harms, overly general deceptive statements require a statement of SOC. However, because the deception is a result of withholding information (rather than making false statements), SOC will have to be modified in this case:

> wSOC: You should be aware that, in order to complete this study, the investigator cannot inform you of all of its details. For this reason, certain details have been left out of the description of the study. However, the investigator will be happy to explain these details to you at the end of your participation (end of the study). In addition, you are free to choose not to participate if you do not like this use of deception, or for any other reason, and your refusal will not be held against you in any way.

With SOC and wSOC in place, I next offer four reasons why SOC should be included in any account of acceptable deception. Because we have already considered most of these issues, my points here will be relatively brief.

a. We have seen that deception per se presents a risk to subjects; at least some subjects are harmed when they are deceived, and other subjects may have relevant idiosyncratic concerns that are never revealed because of the deception. SOC warns subjects of these risks. Therefore, potential subjects who are bothered by deception per se will be able to avoid deceptive studies, subjects with idiosyncratic concerns will have the opportunity to reveal them, and other subjects will have the opportunity to consent to being deceived. SOC also removes any harm to the research team that is incurred when members of the team are required to deceive subjects without their consent.
b. SOC makes our account of acceptable deception internally consistent by ensuring that the use of deception does not violate the second condition (i.e., does not conceal any aspects of the study that might be relevant to subjects' enrollment decisions).
c. SOC allows us to block the undermining of subjects', and the public's, trust in science and medicine, by flagging deceptive studies.
d. I argued earlier that any consequentialist account of acceptable deception—any account that justifies subject deception on the basis of potential benefit to others—faces two major objections: First,

> many people, including many subjects, regard consequentialist justifications as per se unacceptable. Second, developing an accurate estimate of the harm caused by deception, and comparing that harm to the benefits of the study, will be especially difficult. SOC allows us to avoid both of these problems. It allows us to justify exposing subjects to deception, not in terms of the potential benefit to others, but in terms of the subjects' consenting to that deception and, in the process, eliminates the need for a consequentialist justification.

With these points in place, I complete the argument for SOC by considering what I take to be the most important arguments against my account of acceptable deception.

Possible Objections

One might object to SOC for two general reasons: it does not do enough to remove the harm of deception; it does too much to remove the harm of deception. Consider the former objection first. One might argue that subjects cannot provide informed consent for a deceptive study even when they are informed of, and consent to, the use of the deception. In order to provide fully informed consent, subjects would have to know the nature of the deception as well. They would have to know what they are being deceived about. However, because SOC does not require this, it therefore does not provide ethically valid consent. We can grant that subjects cannot provide fully informed consent as long as they are unaware of the nature of the deception, and as long as we understand "fully" in a literal sense. However, as is widely acknowledged, subjects never know everything there is to know about any study. There is simply too much to know. Therefore, in order to make this argument, one will have to show that the information that subjects fail to obtain as the result of deception casts doubt on their informed consent in a way that the other information they never receive does not.

Recall that the second condition states that subjects must never be deceived about "aspects of a study that would affect their willingness to participate. . . ." Roughly speaking, all information must pass this test before it can be omitted from an informed consent. Subjects must be given all of the information that is relevant to their enrollment (or con-

tinued participation) decisions—whether the study uses deception or not. And because information that is withheld as a result of deception must first pass this same test of relevance, I conclude that the two practices are ethically equivalent. One might respond that I have missed the point: deception is different. In cases of deception, subjects fail to have certain information, not because there is too much information to convey, but because they are being deceived. For instance, if Cherek's study had been *non*deceptive, his subjects might have been unaware of how they were losing points simply because no one told them. This information might have passed the irrelevance test. In the actual study, however, Cherek's subjects end up with a *false* belief: that another subject is stealing their points.

Clearly, there is this difference between the two cases, but it is not equally clear that this difference is an ethical one. First of all, we have already seen that SOC informs subjects of the use of deception and provides them with the opportunity of consenting to its use. Therefore, the deception per se will not make an ethical difference between the two cases. The subject, not the researcher, is in control of deciding whether to participate in a deceptive study. In addition, the fact that subjects end up with false beliefs (e.g., "another subject is stealing my points") rather than no beliefs at all (e.g., no belief about how their points are being lost) does not seem to make an ethical difference either. The mere fact that I have some false beliefs concerning X, while a second person is agnostic concerning X, does not, *in itself*, make me any less capable of giving ethically valid consent. It all depends on the X: is the issue in question relevant to providing informed consent? And, as we have seen, relevance to informed consent is decided using the same method in both cases: is the information relevant to the subject's decision to enroll? If the answer is no, then, in both cases, withholding that information does not cast doubt on the subject's informed consent.

There is one final possible difference: subjects of deceptive studies may develop positively harmful beliefs as the result of being deceived. For instance, subjects who are deceived into believing that they are taking a toxic drug may very well be harmed as a result of having that *belief*, even though they are taking a placebo and, thus, presumably, will not be harmed by the capsules themselves. This is a risk of deceptive studies. However, what subjects are told in the process of being deceived is an explicit part of deceptive studies and can be evaluated by IRBs in the same way that the risks of nondeceptive studies are evaluated. Also,

in these cases, the possible harm traces to the fact that subjects believe they are doing something worse than they are in fact doing. Therefore, if anything, the net result of this risk will be to skew subjects' risk/benefit analysis away from, rather than in favor of, research participation.

Along somewhat different lines, one may worry that if we use SOC to justify relatively benign deception today, we may end up using it to justify seriously harmful deception tomorrow. Fortunately, SOC offers several ways of blocking this descent toward Tuskegee. First, the neutral risk condition explicitly prohibits harmful deception: subjects cannot be deceived about any aspects of the study, such as risks or discomforts, that might affect their decision to enroll. Second, researchers presumably will be reluctant to carry out seriously harmful deception as long as they are required to tell subjects about it afterward. Therefore, including the requirement of mandatory debriefing should help minimize the risk of this slippery slope. Finally, I suspect the worry that adopting a policy on acceptable deception will lead to subject abuse is based largely on the belief that deceptive studies will become too easy to conduct. Without the safeguards of informed consent, the cost/benefit calculation inherent to research may become skewed away from present subjects toward future medical progress. Less abstractly, if we allow researchers to lie, they may end up doing things to subjects that, in an honest moment, they would not consider doing. In fact, SOC should actually discourage the use of deceptive studies because it introduces an additional variable into deceptive studies—the variable of how subjects' knowing that they are being deceived will affect their behavior. Therefore, when faced with this requirement, researchers will presumably first look for nondeceptive methods that avoid introducing this additional variable.

To come to the second objection, the fact that SOC introduces an additional variable may lead some to conclude that it requires too much of a sacrifice to avoid the harm of deception: the introduction of an additional variable may in itself confound the results of deceptive studies. For instance, subject behavior may be affected if subjects are apprised of the use of deception and then spend the entire study trying to discover its source. Although this is a possibility, there are several ways in which this effect can be reduced. First, the present account requires that researchers make a clear offer of debriefing. Knowing ahead of time that they can eventually learn the nature of the deception should reduce subjects' desire to discover it for themselves. In addition, by admitting the use of the deception, researchers will be able to enlist subject cooper-

ation in maintaining the deception. For instance, subjects can be asked ahead of time to focus on the procedures specifically asked of them, and researchers can explicitly limit any probing questions.

Finally, SOC does not introduce an extra variable as much as it replaces one variable with a slightly different one. It is fairly common knowledge (particularly among college undergraduates who participate in many of the deceptive studies) that some studies are deceptive and that certain sorts of studies tend to be deceptive more often than others. For this reason, the confound that is introduced into a study when subjects spend their time trying to locate the source of deception is already present to some extent. Thus, although SOC makes the use of deception explicit and, hence, increases the number of subjects who are aware of its use, this openness also provides the opportunity to address the issue directly with subjects. The overall effect may well be a reduction in the degree to which the use of deception affects the results of research.

One might next argue that SOC will reduce subject accrual. Now although this is a common sentiment, there is, as far as I know, no evidence to support it. On the other hand, we have seen clear evidence that deception is harmful. Therefore, given that the present approach *may* eliminate these harms, without introducing any harms of its own, one cannot reject SOC by citing the fact that it might reduce subject accrual. Instead, as in all other areas of science and medicine, we will need to do the necessary studies before drawing any final conclusions. Conducting the appropriate studies would also allow researchers to fine-tune SOC. For instance, it might turn out that subject accrual is reduced, and subjects still feel harmed at debriefing, as long as the statement of SOC is very general: for instance, stating simply that deception is being used. Some subjects may be scared off by such a statement, and those who participate may still be upset when they learn exactly how they were being deceived. If this turns out to be the case, researchers could specify which aspect of the study (e.g., the purpose, a particular procedure) is being misdescribed (or withheld). In this way, subjects whose moral concerns related to particular purposes could participate (and not be surprised) when the deception involved a specific *procedure*, and subjects who have concerns about doing certain things (e.g., playing computer games) could participate on being informed that the deception involved the *purpose* of the study. Finally, even if it turns out that SOC seriously affects subject accrual, it does not necessarily follow that SOC should be rejected. The fact that SOC substantially decreased subject accrual would

suggest that subjects are especially bothered by the use of deception. Hence, we would have even more reason to argue that the use of deception *should* be revealed during the consent process (even if doing so would require that more potential subjects be considered for enrollment).

My response here assumes that subjects accurately gauge the extent to which being deceived will bother them. However, SOC will present problems if this is not the case: if, for instance, *awareness* of the use of deception concerns subjects far out of proportion to how much being deceived bothers them. In that case, SOC will make research less effective, without making it significantly less harmful. In general, subjects would be upset more by the *prospect* than by the use of deception if they thought that the deception might be concealing some risks involved in the study. In this regard, I argued earlier that including a statement that the deception does not conceal any risks might needlessly raise subject concerns about nonexistent risks. I pointed out that for just this reason we do not list all the possible side effects of a particular drug and then tack on the claim that there are no additional side effects. With respect to *non*deceptive studies, doing so would presumably only raise unwarranted concerns about nonexistent risks. One could argue, however, that this result depends upon subjects' initial assumption that researchers are honest. Subjects go into a study assuming that the listed risks exhaust the list of risks. Subjects may dismiss this assumption and, in the process, develop fears of concealed risks once they learn of the use of the deception. The solution (to the extent that this turns out to be a problem) would be to state explicitly that the deception involved in the study does not conceal any risks. In other words, contrary to what I stated earlier, it might be helpful to include a statement that the neutral risk condition has been met. In the end, which alternative we should adopt here depends upon how subjects react to SOC. Thus, a final decision will have to await the appropriate studies called for previously. My argument has been that conducting such studies, and thus developing the optimal approach to SOC, should be a project of high priority. Our ability to reconcile respect for subjects with the occasional scientific need for deceptive research depends upon it.

Summary

I have argued for a concrete policy on when no expected benefit subject deception is acceptable. I started this argument by examining the Ethical

Principles of Psychologists and Code of Conduct (the "principles"), a notable exception to the lack of specific conditions on acceptable deception. After making several minor revisions, I argued that the principles allow deception in research as long as it passes two tests:

1. The extent of deception is minimized (as determined by the second half of condition 1, and by conditions 2 and 3).
2. Any remaining deception is justified by the study's potential benefit to others (as defined by the first half of condition 1).

I then argued that, understood in this light, the principles suffer from two serious flaws:

1. Consequentialist justifications are widely condemned within the context of medical and behavioral research.
2. This consequentialist approach will be almost impossible to implement because there is no way either to assess accurately the harm of deception or to compare that estimate to the potential value of the study.

In light of these two difficulties, I argued that a more satisfactory approach would be to require that subjects be informed of the use of deception *prior* to enrolling in a deceptive study. I pointed out that "second order" consent, or SOC, has two versions, depending upon whether the deception is the result of misdescription or the withholding of information. I defended the SOC approach to acceptable deception on several grounds:

1. SOC justifies the use of deception on the basis of subject consent, thus allowing us to avoid both the contentiousness of consequentialist justifications in general and the difficulties with their implementation in particular. Requiring that subjects consent to the use of deception also restores subject autonomy by permitting subjects to control whether or not they are deceived.
2. Some percentage of subjects are harmed by deception per se. Therefore, simply ensuring that the deception does not conceal any risks will not serve to remove all of the hidden risks from deceptive studies. SOC makes the risks associated with participating in a deceptive study explicit.

3. SOC removes the possibility that the use of some deceptive studies will undermine subject trust in nondeceptive studies.
4. SOC removes the possibility that the use of some deception will undermine public trust in science and medicine.
5. SOC removes much of the harm that running a deceptive study places on the investigative team. Researchers will no longer be guilty of deceiving subjects without their consent during the study, and they will not have to face the difficulties involved in debriefing subjects of that fact once the study has ended.

I finished by considering two possible objections to SOC. First, I argued that it is consistent with ethically valid consent. Even though subjects do not know the nature of the deception, they are aware both of the use of deception and (if the satisfaction of the neutral risk condition is explicitly stated) of the fact that the deception does not conceal any risks. Therefore, from an ethical perspective, these subjects are in the same position as subjects of nondeceptive studies. No one knows everything, but they have all been informed of the aspects of the study that are relevant to their decision to enroll. I next argued that SOC does not make excessive demands on researchers:

1. SOC should not seriously affect subject accrual. However, if it does, that fact itself suggests a need for this policy.
2. I granted that SOC may introduce a (different) variable into deceptive studies. However, this effect will be constant across all research subgroups and can be minimized by dealing with the issue ahead of time and emphasizing the possibility of debriefing.

Finally, I pointed out that the argments I have presented are theoretical ones. However, they offer the possibility of empirical testing. Therefore, the next step is to stop debating at a theoretical level. With second-order consent in place, we have the opportunity to test whether or not deceptive research can be consistent with respect for research subjects.

References

American Psychological Association. 1992. Ethical Principles of Psychologists and Code of Conduct. *American Psychologist* 47:1597–611.

Cherek, D.R. 1990. Human Aggressive Response Maintained by Avoidance or Escape from Point Loss. *Journal of Experimental Analysis of Behavior* 53:293–303.

Fleming, M. 1989. Informed Consent, Deception, and the Use of Disguised Alcohol Questionnaires. *American Journal of Drug and Alcohol Abuse* 15:309–19.

Fost, N. 1975. A Surrogate System for Informed Consent. *Journal of the American Medical Association* 233:800–3.

Office for the Protection from Research Risks. National Institutes of Health. *Protecting Human Research Subjects: Institutional Review Board Guidebook*. Bethesda, Md.: National Institutes of Health.

Additional Readings

Asch, S.E. 1952. *Social Psychology* . New York: Prentice-Hall.

Baumrind, Diana. 1964. Some Thoughts on Ethics of Research: After Reading Milgram's "Behavioral Study of Obedience." *American Psychologist* 19:421–3.

Beauchamp, T., and J. Childress. 1989. *Principles in Biomedical Ethics*. New York: Oxford University Press.

Bok, S. 1978. *Lying: Moral Choice in Public and Private Life*. New York: Random House.

———. 1992. Informed Consent in Tests of Patient Reliability. *Journal of the American Medical Association* 267:1118–19.

———. 1995. Shading the Truth in Seeking Informed Consent for Research Purposes. *Kennedy Institute of Ethics Journal* 5:1–17.

Cherek, D.R. 1993. Acute Effects of Marijuana Smoking on Aggressive Escape and Point-Maintained Responding of Male Drug Users. *Psychopharmacology* 111:163–8.

Fost, N. 1979. Consent as a Barrier to Research. *New England Journal of Medicine* 300:1272–3.

Freedman, B. 1982. The Validity of Ignorant Consent to Medical Research. *IRB: A Review of Human Subjects Research* 4:1–5.

Milgram, S. 1963. Behavioral Study of Obedience. *Journal of Abnormal Psychology* 67:371–8.

———. 1964. Issues in the Study of Obedience: A Reply to Baumrind. *American Psychologist* 19:848–52.

———. 1977. Subject Reaction: The Neglected Factor in the Ethics of Experimentation. *Hastings Center Report* 7:19–23.

Patten, S.C. 1977. The Case that Milgram Makes. *Philosophical Review* 86:350–64.

Acknowledgments: This paper was first presented at a National Institutes of Health, Bioethics Program colloquium. I extend my thanks to the participants for their input, to John Gluck, and to several anonymous reviewers for their helpful suggestions.

Address correspondence to: Dave Wendler, Ph.D., Bioethics Program, Building 10, Room 1C 116, Bethesda, MD 20892-1154.