How (Un)acceptable is Research Involving Deception? by Charles P. Smith

The use of deception in research has sparked considerable controversy. At one extreme are authors who object to deception in principle and who aver that alternative research approaches are feasible (e.g., Diana Baumrind, Margaret Mead, and Thomas Murray). Differing sharply are those who feel that the use of deception is legitimate and that it is necessary in the sense that some problems are not investigable by any other means (e.g., Robert Baron and Stanley Milgram).1

The research to be presented in this article might be considered an empirical experimental approach to research in ethics. I will report a study of participants’ reactions to deception and discuss the implications of the results for the deliberations of IRBs. The term “deception” will be used to refer to either deliberately withholding information or deliberately misleading or misinforming the subject.

Why, one might ask, is deception employed so frequently in some branches of behavioral science research? Primarily, I think, for two different reasons: first, because it is perceived as necessary; that is, because it is believed that the full revelation of the purpose or procedures of the research would vitiate the attempt to study the problem under investigation. Second, and quite distinct, deception is a relatively easy, and often highly effective, way of creating a desired condition, such as a feeling of competition, or the opportunity to help a person in need. To the best of my knowledge, deception is rarely employed in order to trick subjects into a situation they would not knowingly have entered.

The investigator who contemplates the use of deception is faced with the dilemma of how to acquire knowledge without infringing on the rights and welfare of research participants. A deontological approach to this dilemma deals with the matter on the basis of principle (or the weighing of principles). For example, deception might be considered wrong because it violates trust, or because the subject is treated as a means rather than as an end. Empirical considerations are not relevant to the resolution of the dilemma. In contrast, an approach that evaluates alternatives in terms of their consequences asks whether deception is harmful, and whether the benefits outweigh the harm. Empirical evidence is relevant, since whether or not a procedure has harmful or beneficial effects is an empirical question.

It follows that the negative reactions to deception that would be of concern to deontologists should be different from those that would concern consequentialists. Murray, for example, distinguishes between being “wronged” and being “harmed.”2 At p. 8 The deontologist might focus on violation of trust or autonomy, or on being “used”; the consequentialist might focus on harm or on the fact that the participant did not have a worthwhile learning experience. A psychologist can attempt to make a contribution to these issues by assessing both kinds of reactions to deception.

Reactions to Deception

To begin, I shall describe some research I conducted in collaboration with Susanne Berard and Carl Malinowski.3 Earlier Malinowski had studied the relationship between Kohlberg’s stages of moral reasoning and cheating on the part of college males.4 Kohlberg has identified qualitative changes in reasoning about moral dilemmas that occur as a person matures. Most college students are found to reason at a “conventional” level (which includes a lower and a higher stage of reasoning). Malinowski found that students at the lower stage cheated substantially more than those at the higher stage. Participants were strongly tempted to cheat in his achievement-oriented testing situation in order to avoid embarrassingly low scores, and, in fact, 41 of the 53 subjects did cheat by reporting falsely high scores on a task involving motor skill.

Considerable deception was employed. First the subject was told that the purpose of the experiment was to study attention and concentration when the real purpose was to study...

Charles P. Smith is Professor of Psychology, Graduate School and University Center, City University of New York.
cheating. Second, false norms were provided to make the subject feel that his performance was lower than that of other students. Third, the subject was given the impression that his performance was not being observed, whereas it was being recorded in the adjoining room. Finally, some subjects, before being tested, talked to "the previous subject" (a confederate) who implied that it was necessary to cheat in order to obtain a decent score.

Whatever you may think of this elaborate chain of deceit, it was not gratuitous. Each element served a theoretically relevant purpose in establishing the conditions necessary for testing the hypothesis. However, Malinowski's IRB felt that his proposal presented problems of deception, informed consent, invasion of privacy, stress, and entrapment. The proposal was approved only after careful procedures were worked out to ensure anonymity and to provide a thorough debriefing. After the debriefing subjects were given an opportunity to take their data and withdraw or to turn in the data and sign a consent form. As it turned out, every subject agreed to have the data used and none complained subsequently.

About a year later I proposed some follow-up research on cheating using a similar method and including the safeguards adopted by Malinowski. Even with the safeguards, my IRB (an entirely different group) raised similar objections to the procedure. They felt it might have adverse effects on the subjects. At this point I decided to try to document the effects on the subjects of participation in Malinowski's cheating study, so my colleagues and I designed a questionnaire asking about various ethically relevant aspects of the research. The respondents were to reply anonymously.

Since nearly two years had passed, it was difficult to track down the original subjects, but we managed to obtain completed questionnaires from 36 of the 44 for whom we could get current addresses. (It does not seem likely that nonrespondents would have been more critical than respondents, since the questionnaire provided a ready-made opportunity to express negative attitudes.) Questionnaires were also sent to two comparison groups of subjects who had participated in studies that posed few, if any, ethical problems. The method and results are reported in detail in the article by Smith, Berard, and Malinowski.

Both in absolute and in relative terms there were few complaints on the part of Malinowski's subjects; 97% felt that their participation had been voluntary and none felt that his identity would not be protected. Three-quarters said that the study had involved deception, but not a single one of those felt the deception had been harmful. Finally, the overall response to having participated was either positive or neutral for all but one of Malinowski's subjects. That one checked "negative" rather than "very negative."

While these results do not constitute a license to deceive, they do suggest, in retrospect, that the concerns of the second IRB were exaggerated. The reactions of nearly all of Malinowski's subjects were, if anything, more favorable to the cheating study than were the reactions of the other subjects to the comparison studies.

Hidden in this way of reporting the results, however, is what might be called the problem of extreme cases. That is, even if the overwhelming majority did not have a bad experience, what about the few subjects who considered some aspect of the research experience to be negative? For example, even though the average anxiety, embarrassment, and guilt reported by the cheating study subjects was between "slight" and "moderate," a few persons reported that they experienced "quite a lot" of these feelings. Is the procedure justified if even a few subjects experienced feelings of this sort? Perhaps part of the answer is given by the subsequent ratings by these subjects of their overall response to participation. None of the subjects who experienced these relatively intense feelings gave an overall negative response.

Another five of the cheating study subjects indicated that they felt the procedures entailed ethical problems. Only one of these had a negative overall response to participation.

Part of the solution to understanding the significance of extreme responses is provided by information from a control or comparison group. There may be a few individuals in every large group who are indiscriminately angry and critical. We might call them "soreheads." For example, if 5% of the subjects complain strenuously even about innocuous research, then some of the complaints of the "extreme cases" in the cheating study may be attributable to the unreasonableness of the persons rather than to the unreasonableness of the procedures. In our study the proportion of negative responses in the comparison groups provides a baseline for comparison. As it turned out, the percentage of subjects whose overall reaction to participating was negative was slightly less in the cheating study (2.8%) than in the combined comparison groups (5.9%).

A problem still remains, however, because even if some subjects are indiscriminately negative others may have complaints that are specific to the research in question. For example, if a subject says: "My reaction is negative because I felt embarrassed, or because the experimenter lied to me," then we cannot write him off as a sorehead.

Should further research of this type be approved even if even a very few subjects are likely to find it harmful or morally objectionable? It does not seem appropriate to make such a decision in the abstract. In the case of harm, for example, it would be necessary to know how much harm was involved and the potential importance of the research. Alternatively, if a few subjects say the research is morally objectionable and raise ethical considerations that seem compelling to an IRB, then the research should be discontinued. The best solution, of course, would be to find a way of screening out in advance persons who object to the use, in research, of deception, invasion of privacy, etc. One such screening method has been developed by Mannucci.

Possible Cognitive Dissonance

Now let us return to the generally exculpatory implications of the responses of the cheating study subjects. They are reminiscent of the responses of the subjects in Milgram's study of obedience in which only 1.3% were in the negative range. Are subjects really so positive about research that puts so many of them in a bad light, or are their post-participation evaluations suspect? Could they, for example, be defending against feelings too painful to acknowledge, such as guilt or of being duped? Perhaps they come to believe that if they have voluntarily endured a negative experience it must have been for some good purpose. If the debriefing assures them that the research is of value, that may help them reconcile themselves to the experience. If a subject has behaved in a way that is discrepant from his beliefs, perhaps he will attempt to reduce "cognitive dissonance" and a feeling of being "used" by coming to believe that the experience was valuable and interesting.

Susanne Berard and I investigated this possibility by obtaining reactions to a description of the cheating study from introductory psychology students. We met with them during class time in order to ensure that all students would participate and not just a
self-selected group of volunteers. A detailed description of the cheating study was given by a graduate student who played the role of the experimenter and used Malinowski's script verbatim, with the exception of a few deletions and a few transitional statements. Then, imagining that they had participated as subjects, the students anonymously completed the same questionnaire that the original subjects had completed. Since the responses of the 27 males and 41 females were essentially similar, they were combined and compared with those of Malinowski's subjects.

The results are relatively clear. On the whole the role-taking subjects differ little from the original participants. On no question that specifically asked about ethical problems was there a statistically significant difference between the two groups. However, compared to the real subjects, the role-taking subjects found the experiment significantly less interesting, and they were significantly less positive in their overall reaction. (Most of their responses were in the neutral range, but 7% were in the negative range.)

These results are consistent with the possibility that the cheating study respondents reported greater interest and a more positive reaction as a way of resolving cognitive dissonance. We regard the results as ambiguous, however, because the cheating study was indeed probably less interesting to someone who merely listened to a lengthy description of it than to someone who actually participated. The real subjects were undoubtedly more involved, had a greater likelihood of learning something of value, and felt the debriefing was more personal and meaningful. Although the data do not provide definitive evidence concerning the possible effects of dissonance, they do indicate that ethical issues were no more troubling for the role-takers than they were for the real subjects.

Responses to a question on deception are worth special comment. The question was: "Did the research involve deception?" When 23% of the original cheating study subjects said No, we thought perhaps they did not remember, or that they thought it was no longer deception when the experimenter revealed the true purpose of the experiment. We were especially surprised, therefore, when 43% of the role-taking subjects said there was no deception. We first thought that these subjects simply had not been paying attention, and, while that may have been true for some, that interpretation does not seem consistent with other aspects of the data. After discussing the matter with some of the students we have concluded that for some "deception" was too strong a term. This interpretation is consistent with that independently arrived at in Mannucci's research.

Understanding Negative Responses

On the whole the results from the role-taking group reveal little concern about the ethically problematic aspects of the research, suggesting once again that the concerns of the second IRB were exaggerated. However, there were several extreme cases and their responses are of interest. Of the 67 role-takers, three reported an overall response to having participated as "negative" and two more checked "very negative." One commented: "We were not told what the purpose of the experiment was, therefore there is no value because not knowing we felt (a) anxiety, (b) skeptical, (c) embarrassed, (d) used. There has to be a way of testing a human being with an understanding and with that understanding a willingness on his part." Three of the five subjects checked that they felt their (imagined) participation had been "somewhat coerced." A feeling of coercion was a major determinant of a negative overall response in the role-taking group (r = .47, N = 67).

Although these negative responses seem quite clear and specific, their interpretation is by no means straightforward. These criticisms may be thoughtful and warranted, or they may stem from irrelevant sources. Four such sources are suggested by the questionnaire responses: (1) Difficulty in role taking is suggested by responses that confused the procedures used to obtain the role-taking data with those used in the original cheating study. For example, there was little, if any, coercion in the original study. (2) Inattention to the description is suggested by incorrect answers to factual questions. (3) Indiscriminate negativism was reflected in the responses of some who checked "no" or "don't remember" for nearly all the items. These may have been "soreheads." (4) A negative attitude toward social science research may have prevented some persons from being open-minded to the potential value of the cheating study. The one negative subject in the original cheating study may exemplify this form of bias. He comments: "I don't feel psych research has any scientific purpose . . . It has no significance as a science . . . Science is a reality and psych is not." Thus, it is difficult to determine whether a negative comment is a valid objection or whether it stems from such irrelevant factors as indiscriminate negativism, inattention (and/or misunderstanding), or bias against social science research.

I mentioned earlier that deontologists might react differently to deceit from consequentialists. Like most philosophers of ethics, most research participants probably hold a mixture of deontological and consequentialist beliefs. In fact, most of the respondents who complained objected on both types of grounds. However, some seem more deontological than others. For example, subject 309 feels he was somewhat coerced and taken advantage of (violation of autonomy, being used) and he specifically indicates that while he does not regard the deception as harmful, he does not feel it was justified by the scientific purpose of the research. He seems to be objecting to deception on principle rather than because of its consequences. Of the subjects who said that the deception was harmful, some said it was justified and others said it was not. Perhaps some of these subjects were consequentialists. Unfortunately, because of the way the questions were worded and the paucity of negative comments this distinction does not emerge clearly from the present data. It should be possible, however, to pursue it more effectively in future research.

Another matter of interest is why both the real subjects and the role-takers were apparently less concerned about the ethically problematic aspects of the cheating research than the IRB members. Five possible reasons come to mind:

1. The subjects were better informed about the research in question than the IRB members. The subjects experienced the research in its full detail whereas the IRB members only read a proposal describing the research. The proposal may not have provided sufficient information to permit an accurate judgment.

2. As a general rule subjects appear to be less concerned about deception and invasion of privacy than IRB members. Several studies have shown that subjects regard deception as only mildly objectionable and that they view it as a necessary and justifiable aspect of research. Mannucci's research indicates that what matters most to subjects are the consequences of deception (e.g., stress, humiliation), not the deception per se.

3. IRB members have accepted the role of protector of research participants. Hence they are likely to be especially sensitive to possible problems
and to err on the side of caution. They may also be more knowledgeable about ethical issues than subjects. In addition, the consequences of making judgments about deception are not the same for subjects and for IRB members.

4. As they read research proposals IRB members may give labels to various aspects of the procedures. For example, they may think "This is a cover story." "This is deception." "This is invasion of privacy." A process called verbal mediation or mediated generalization may then occur. The reader may then react, not to the original description, but to the assigned label with all its negative connotations.

5. When the proposed methods have been used before, the judgments of the IRB members may be affected by knowledge of the results of these methods. Milgram has pointed out that there were no objections to his obedience procedure until the effects of the procedure were known. In the case of the cheating research, the fact that 41 of the 53 subjects in the original study cheated may have led the IRB members to infer that the situation was too stressful or that the subject was entrapped. While it is entirely appropriate for an IRB to evaluate a procedure on the basis of preliminary results, it may not be appropriate to make evaluations based on inferences from results (e.g., many subjects cheated, therefore it must have been stressful).

Is Deception Research Permissible?

To conclude, let us return to the question of whether research involving deception should be permitted. I suggested that a psychologist could help to answer this question by assessing the consequences of deception. As we have seen, this turns out not to be as simple as it sounds since both the positive and negative reactions of subjects may be biased, at least to some extent under some circumstances. Indeed, the direct verbal reports of subjects may need to be supplemented with indirect measures such as willingness to volunteer again for research. One may conclude, however, that the procedures employed in the cheating study produced very few negative reactions since the verbal reports of the subjects were consistent with the fact that there had been no spontaneous complaints about the research and with the reactions of a role-taking group.

Obviously one cannot generalize from the cheating study to all research involving deception, but with further improvements such assessments can help to establish whether various procedures cause harm or are otherwise unacceptable to research participants. Such information should help IRBs to evaluate the risks of proposed procedures more accurately.

Systematic research remains to be done on such topics as the effects of deception on the experimenter, the effect on graduate students of the deceptive experimenter as a model, and the effects on the profession and on society of deceptive research practices. Certainly one consequence already manifests is that subjects are often suspicious even when studies involve no deception.

Most research participants, at least those who are college students, seem to have fewer reservations about deception than IRB members. Perhaps this is the case because they have been made callous by the amount of deception they encounter in everyday life, but I think there is another more appropriate reason, namely, that deception per se has a different meaning in the context of scientific research than it has in everyday life. Most subjects understand that deception is a necessary aspect of some kinds of research and find it justifiable under such circumstances.

The view I am presenting here is different from that of Milgram who has suggested that an experiment is like a theatrical performance or a magic show. On those occasions members of the audience know they are being deceived. There is no problem of informed consent. In a deceptive experiment the subject cannot know until afterwards that he was deceived, and he can only consent to have his data used after the deception has occurred.

When is deception a problem then? Clearly it is problematic when it has harmful consequences or when it is used to trick a subject into a situation that he or she would not have entered voluntarily. But does deception always come into conflict with the principle of informed consent? It seems to me that it does not. The purpose of assuring informed consent is not to give a subject complete information about an experiment, but to give whatever information might be expected to affect the decision to participate. Deception may or may not conceal factors that the potential subject would want to consider in making that decision. Problems that (apparently) cannot be studied without the use of deception entail practices that range along a continuum of acceptability to subjects. At one end are those to which most participants would have no objection (e.g., not informing the subject that his speech pauses are being studied). At the other end are those that entail unpleasant experiences, unwelcome revelations about oneself, or the scrutiny of behavior that the subject ordinarily tries to conceal (e.g., stealing). It seems to me that a problem of informed consent occurs only toward the latter end of this continuum.

Our experience in conjunction with the research reviewed above suggests that there are three considerations that help to make deception acceptable to a research participant. First, the consequences must not be too unpleasant. Second, the subject must not feel coerced; he or she must feel that to some extent at least participation was voluntary. Third, he or she must feel the research is justified by its scientific importance. Moreover, to the question of whether deception in research should be permitted is a carefully qualified yes.

ACKNOWLEDGMENT

This paper is adapted from a talk given at the Hastings Center on November 3, 1980. Valuable comments and suggestions were received from Stefan Baumann, Susanne Berard, R. Glen Hass, and Judith Smith. Special thanks are due to Hilary Liberty for his help with the role-taking study.

REFERENCES


7Epstein, Y. M. et al. op. cit.