Scientific Rewards and Conflicts of Ethical Choices in Human Subjects Research

Peter David Blanck
Alan S. Bellack
Ralph L. Rosnow
Mary Jane Rotheram-Borus
Nina R. Schooler

College of Law, University of Iowa
The Medical College of Pennsylvania
Temple University
Columbia University
University of Pittsburgh

The primary responsibility of the American Psychological Association's (APA) Committee on Standards in Research (CSR) is to advise the APA on issues and standards related to the protection of human participants in psychological research. A related goal is to enhance the use of good ethical practices by APA members. The purpose of this article is to foster the view of research ethics not as an affront to the integrity of sound research, but as opportunities for scientific rewards, including increased understanding of the meaning of data, enhanced recruitment, and the inclusion of more representative samples. Three ethical practices are discussed as examples of this general premise: respect for confidentiality, use of debriefing, and assurance that participants are noncoerced volunteers. The Committee's intent is to promote consideration of these issues, not to promulgate specific guidelines or procedures.

The main purpose of the Committee on Standards in Research (CSR) has been to advise the American Psychological Association (APA) on issues and standards related to the protection of human participants in psychological research. In recognition of the increased variety of ethical concerns facing the field, the scope of CSR has broadened considerably. Recent efforts have addressed such topics as informed consent, confidentiality of human participants, and scientific misconduct. The Committee believes these topics are essential in the efforts to enhance the protection of human participants. The suggestions emanate from the different backgrounds and perspectives of the CSR members—laboratory and clinical researchers, social and clinical psychologists, and an attorney.

In the 1970s, in the wake of growing concern over what were seen as moral issues, the APA set out to formulate a code of ethical practices to govern research involving human participants. An ad hoc committee developed a document that was adopted by the Council of Representatives in 1972 and issued as an informative booklet the following year (APA, 1973). The document was revised a decade later by the Committee for the Protection of Human Participants in Research (APA, 1982), and a further revision is currently under way by the CSR. At the core of this document is a set of principles representing ethical practice and linked to the use of stringent safeguards to protect the rights of research participants. In the past 10 years there has also been a proliferation of governmental and institutional regulations that give voice to the participants' rights in human research. Last year, for example, the United States Office of Science and Technology issued regulations that set forth a comprehensive federal policy for the protection of human subjects (Federal Register, 1991, p. 28003). As a consequence of these efforts, psychological researchers have become accustomed to the close scrutiny of institutional and professional review boards as proposals are evaluated in regard to ethical practices. Recent research also suggests that individual researchers differ systematically in the ways they formulate and evaluate ethical principles (Kimmel, 1991).

Since the 1970s, the APA code has underscored the idea that sound research often calls for a delicate balance between humane, moral, and scientific values. For example, it is considered essential that people have freedom of choice in deciding whether or not to participate in psychological research (Sieber & Sorensen, 1991). However, the nature of freedom in choosing to participate is often hazy, and may have a significant effect on the nature of the sample. We will return to this point, but it is recognized generally that, in the absence of a special effort to stimulate participation, the characteristics of persons who volunteer for laboratory and community-based re-

Peter Blanck is chairman of the Committee on Standards in Research (CSR). The other four authors, who are members of CSR, contributed equally to this article and are listed in alphabetical order. Elizabeth Baldwin, APA staff liaison, and Robert Rosenthal provided helpful comments.

Correspondence concerning this article should be addressed to Peter David Blanck, University of Iowa College of Law, Iowa City, IA 52242.
search differ in important ways from those participants with relatively less freedom of choice to participate. This tension is particularly apparent, for example, in research in which persons of lower income levels are induced to participate (Blanck, in press-b; Rosenthal & Rosnow, 1975).

Given the strong relations among methodology, ethics, and research artifacts, the APA code encourages researchers to invest their ingenuity in discovering ways of conducting studies that avoid ethical violations. In this article, we suggest ways in which the APA code can serve also as a window onto opportunities to increase the rewards of ethical choices to researchers who are able to adopt such practices.

We discuss three ethical practices that may engender scientific rewards: (a) the use of confidentiality to protect the privacy of disclosures, (b) the use of debriefing to clarify the nature of the study, and (c) the use of volunteers to assure freedom from coercion to participate. We realize that these practices are not always feasible, such as when the obligation to advance knowledge appropriately over-rides privacy (Blanck, 1987; Blanck & Turner, 1987; Pat-tullo, 1982; Sieber & Sorensen, 1991; Veatch, 1982) or when providing complete information about significant research compromises the validity of the data. Given the “double-edged potentiality” of ethical issues (APA, 1973, p. 8), we suggest some limiting conditions (or potential conflicts) so as not to present an overly optimistic view of what is possible. But the primary objective of this article is to make a number of suggestions designed to foster a view of research ethics not as a hindrance to the integrity of sound research, but as an opportunity for scientific rewards in psychological research with human participants.

Use of Confidentiality

Three basic principles appear in all European and American ethical codes for psychological research: (a) avoid physical harm, (b) avoid psychological harm, and (c) keep the data confidential (Schuler, 1982). The first two principles emanate from the Nuremberg code of 1946–1949, developed in conjunction with expert testimony against Nazi doctors at the Nuremberg Military Tribunal after World War II (Beecher, 1970). Confidentiality, the third principle, evolved to safeguard information divulged by research participants and clients. Confidentiality is commonly justified on the basis of three claims: (a) that researchers have a professional right to keep subjects’ disclosures secret, (b) that fairness requires respect for privacy, and (c) that enhanced credibility or validity should result when the researcher has promised to keep disclosures confidential (Bok, 1978).

The protection of confidentiality can sometimes present legal, methodological, and ethical dilemmas for researchers (Appelbaum & Rosenbaum, 1989). Special problems are apparent and have been documented in detail elsewhere with respect to confidentiality in community-based research projects (e.g., Blanck, in press-b; Sieber & Sorensen, 1991). For example, certain research data (e.g., field studies of child abuse or venereal disease) are ordinarily not immune to subpoena (Knerr & Carroll, 1978; Melton & Gray, 1988; Rozovsky, 1990; Sieber, 1982). Likewise, routine partner notification (as a breach of confidentiality) could limit in certain circumstances the validity of community-based research findings, such as those involving human immunodeficiency virus (HIV) testing (Melton & Gray, 1988; Sieber & Sorensen, 1991). However, in certain sensitive research situations, it is possible to obtain a “certificate of confidentiality” from the Public Health Service (for a review, see Melton & Gray, 1988; Sieber, 1992). These certificates protect participants’ names and other identifying information from being subject to a court’s subpoena power. However, prior research suggests that most ethically sensitive studies involving human participants are not typically covered by certificates of confidentiality (Melton & Gray, 1988).

Most institutions require that consent forms specify that the data will be confidential, but this pledge can be highlighted to different degrees. Assuring participants of the confidentiality of their responses is not simply for their benefit, but may increase the likelihood that they will be honest and open in their responses (Boruch & Cecil, 1979). An experiment by Esposito, Agard, and Rosnow (1984) tested the hypothesis that a written assurance of confidentiality would improve self-disclosures by college-age research participants. The participants were administered Spielberger’s (1979) Trait-State Personality Inventory and Crowne and Marlowe’s (1964) Social Desirability Scale under one of two conditions. In the experimental condition, the instruction page asked for the respondent’s name and contained the statement “Your responses on these measures will be kept strictly confidential.” In the control condition, the instructions asked for the respondent’s name but contained no mention of confidentiality. Participants’ self-ratings of anxiety, curiosity, and anger showed lower correlations with social desirability in the confidentiality than in the control condition. This finding suggests that confidentiality can attenuate evasive answer bias.

Circumstantial support for the view that a written or verbal assurance of confidentiality promotes more honest disclosures comes from a number of other areas, although the pattern of results is not unequivocal. Ceci and Peters (1984) observed that letters of recommendation written by faculty advisors were more critical when the form indicated that the student had waived his or her right to inspect the letter. Likewise, Merluzzi and Brischetto (1983) reported that male undergraduate students evaluated counselors as less trustworthy when the counselors had breached confidentiality. A survey study by Singer (1984) found that the great majority of respondents believed that assurances of confidentiality foster cooperation in answering questions. On the other hand, Reamer (1979) reported no effect of confidentiality in an interview study of youths previously arrested for status offenses.

This topic requires a more detailed analysis because little is known about the general magnitude of the rela-
tionship between verbal and written assurances of confidentiality and the subsequent quality of data in human participant research (cf. Committee on Federal Statistics, 1979). The size of the relationship obtained by Esposito et al. (1984) is modest by conventional standards of psychological research, but is not unimpressive when recast in terms of its practical significance (e.g., Rosenthal, 1990). The effect sizes in Ceci and Peters’s (1984) study seemingly ranged from moderate to substantial, corresponding to different questions that were asked of the faculty advisors. It is still unclear what might be expected of research participants in diverse situations. For example, volunteers for studies of a new medication for acquired immunodeficiency syndrome (AIDS), marital conflict, or sexual practices would likely value confidentiality much more than college students in a study on perceptual acuity. Conversely, parents of child participants in a clinical trial generally demand information about their children, a potential breach of the child’s confidentiality. It would be important to study the effects of verbal and written assurances of confidentiality in such highly sensitive research studies, such as those involving persons at risk or children (Grisso et al., 1991). It is also worth noting situations in which clear guarantees of confidentiality are essential for recruiting representative samples. Research on AIDS is a case in point. In such cases the usual pro forma statement will not be sufficient to allay the concerns of potential participants who are terrified by the possibility of public disclosure.

Thus, researchers face a difficult task to ensure confidentiality in a climate of shifting ethical standards (Bayer, 1985). As suggested above, the AIDS pandemic presents researchers with such challenging ethical dilemmas. Researchers must ensure confidentiality knowing that the rationale guiding the procedures established to protect participants’ disclosures may shift over time (Bayer, 1985; Bayer, Levine, & Murray, 1984; Bayer & Toomey, in press). For example, several large longitudinal research studies examining HIV risk acts (e.g., the Multi-site AIDS Cohort Study, the New York City Gay Men Study) were ongoing when HIV testing procedures were developed. Because of the particular data collection procedures, it was possible for researchers to know the HIV serostatus of the participants and for participants to choose to know their status or not. Researchers had the opportunity to examine the effect that knowledge of one’s HIV serostatus had on behavior change and on participants’ reported quality of life (Ostrow, 1991; Martin, 1987). Participants who were tested for HIV could choose not to be informed of their serostatus. Later, the National Institutes of Health adopted a policy that it was unethical to conduct research in which the participants were tested for HIV and not informed of their serostatus. Participants who chose to get tested for HIV and declined to be informed of their serostatus under earlier guidelines no longer had the option to remain uninformed. The circumstances existing at the time of choosing to participate shifted during the duration of the study. The AIDS pandemic has required shifting community standards (Dickens, 1988). Similar concerns exist around the issues of confidential versus anonymous testing (Annas, 1988; Curran, Gostin, & Clark, 1988), partner notification (Bayer & Toomey, in press), recruitment of pregnant users of intravenous drugs whose children may be born addicted (Macklin, 1990), and HIV testing of adolescents (Rotheram-Borus, 1991b).

Use of Debriefing

The word debriefing has its roots in military jargon; it was first used during World War II to refer to the process of interrogating pilots who had returned from bombing missions. In its current usage in psychological research, the term emphasizes a kind of catharsis after treatment. The purpose of debriefing is to remove any misconceptions and anxieties that the participants have about the research and to leave them with a sense of dignity, knowledge, and a perception of time not wasted (Harris, 1988). Jones and Gerard (1967) suggested that debriefing should regularly include discovering what each participant thought of the research situation, thereby providing a more personal context in which to interpret the nature of the results.

In cases in which deception or misdirection is used as part of an experimental design, debriefing is also meant to remove any “detrimental impact on the participant’s feeling of trust in interpersonal relationships” (APA, 1973, p. 77). It might be hypothesized that the revelation that a deception was part of the study could Spawn skepticism and suspicion, which in turn could influence future behavior (either as part of this or another study or in other life activities). However, there is little evidence of any changes in behavior when debriefed participants, even those previously suspicious, participate in subsequent tests or experiments (e.g., Brock & Becker, 1966; Fillenbaum, 1966; McGuire, 1969).

Clinical subjects often participate in research to learn something about themselves, and debriefing is an opportunity to receive feedback regarding performance or response (Sieber, 1992). Participation in clinical research provides a chance to be altruistic by contributing to science and improving clinical services, helping others through participation. Thus, information about one’s own performance or the findings of the study enhances the

---

1 One statistical procedure worth mentioning that has been used to protect participant confidentiality is the randomized response technique (RRT), pioneered by Warner (1965, 1971) for use in large-scale survey research. The participant uses a randomizing device (such as flipping a coin) to select how to respond to a sensitive question (Fidler & Kleinhecht, 1977; Krotki & Fox, 1974). Suppose the question asks, “Have you ever used cocaine?” The subject is instructed to flip the coin, out of the researcher’s sight, and to respond “yes” if it lands heads and “no” if it lands tails. Knowing that 50 percent of the participants are expected to get heads and 50 percent to get tails, it is then possible to estimate the proportion of participants that actually responded that they had sampled cocaine. Although RRT calls for a larger number of participants to produce reliable estimates than may be feasible in most psychological research, the results of RRT might serve as a yardstick for appraising the general effect sizes of simple written and verbal assurances of confidentiality on the quality of research findings.
personal sense of participating and contributing to an important program (Jones & Gerard, 1967).

Debriefing also offers the researcher the opportunity to discover the personal meaning of the study for the participants—whether the experimental procedure was actually perceived or experienced by the participants as the researcher intended. Such information is essential for accurate interpretation of findings. For example, if participants in a pseudo-treatment group guess that their treatment was not real, the value of this condition as a control for positive treatment expectancies is compromised. Similarly, investigators who compare clinical and non-clinical populations may be particularly rewarded by appropriate debriefing strategies, as the meaning of an experimental manipulation or the value of participant payment often differs between groups.

Debriefing has been critical to examining the efficacy of clinical trials for drugs aimed at slowing disease progression for persons infected with HIV. Middle-class gay men in AIDS epicenters (e.g., San Francisco, Chicago, and New York) are well informed about drugs being used to treat symptomatic persons (Gorman et al., 1991). It has been a common practice for participants in clinical trials to share medication with each other, gain access to drugs or treatments available outside of the United States, and take multiple drugs simultaneously. In the supportive, cohesive community climate, it is almost impossible to conduct an evaluation of a single drug uncontaminated by auxiliary treatments. Rather than force research participants to be duplicitous about their multiple treatments and contaminate the protocol in unknown ways, debriefing is a central tool to monitor the degree and type of multiple drug use among gay men in research trials.

Debriefing often can provide researchers with leads for future research and help identify problems in their current protocols. Rotheram-Borus, Koopman, and Bradley (1989) found that adolescents who had received an AIDS prevention program reported during debriefing that they frequently engaged in group sex (i.e., sexual intercourse with multiple partners during a single encounter). Despite substantial pilot work, including focus groups with youths and their counselors, extensive interviews and self-reports of sexual activities, and participation in 10 intervention sessions, these adolescents had not reported their group encounters until the debriefing. When the protocol was subsequently revised to include assessment of sexual encounters with multiple partners, many adolescents conveyed their participation in such activities.

In long-term clinical studies, debriefing is sometimes so much delayed for practical and methodological reasons that when it does take place it is irrelevant to the investigator and the participants may no longer be accessible or interested. Delay may result from the investigator's need not to compromise the short-term conduct of the study or to sacrifice statistical power by interim analysis of an experiment. However, it is usually possible to conduct ongoing post-study interviews with individual participants that allow them to describe their perceptions of the study, even if the debriefing cannot provide complete information about the outcome of the research. More study is needed of the use of post-study debriefing as a source of data in clinical studies. CSR is not aware of a clinical study that has investigated participant satisfaction by comparing a group that was debriefed with one that was not.

Use of Volunteers

We mentioned previously that participant selection bias can upset the balance between methodological requirements and ethics, producing artifacts in the data (Rosnow, in press; Saks & Rosnow, 1981). Artifact refers generally to research findings resulting from factors other than those intended by the investigator (Rosnow, in press). The term does not refer simply to serendipitous findings, but to scientific observations resulting from unrecognized factors that might jeopardize the validity of the investigator's conclusions (Rosnow, in press). For example, to reduce a threat to the external validity of research findings resulting from limited participant samples (e.g., only volunteer subjects), methods for enhancing the diversity of participants are often warranted. The associated ethical issue concerns the point at which the researcher threatens the individual's right and freedom not to participate, such as by offering inducements to participate that are coercive (Blanck, in press-a, in press-c).

In one representative example, Strohmetz, Alterman, and Walter (1990) examined baseline differences in problem severity among alcoholics who did and did not volunteer to participate in a treatment outcome study. The level of the patient's volunteer status (i.e., willingness to participate in the treatment intervention) was positively related to the severity of alcoholism problems reported during the pretreatment period. Although the results can be interpreted in different ways, one plausible implication raised by Strohmetz et al. is that patients who agree to randomization in intervention experiments may somehow be different from the population of interest. King and King (1991) have noted a similar concern regarding intervention research on the adjustment of Vietnam veterans.

As suggested above, ethical practice requires researchers to respect individuals' freedom to decline to participate. However, a number of research strategies have been described to deal with the potential costs of subject selection bias. Rosenthal and Rosnow (1975, 1991) described how, in certain psychological studies in which the population can be stratified into respondents and non-respondents, it is possible to assess the direction of subject selection bias (cf. Saks & Blanck, 1992). They also noted a number of empirically derived strategies for improving the representativeness of subject samples by inducing more nonvolunteers to enter the sampling pool. For example, volunteering rates are likely to increase the greater the material incentive to participate and the less aversive, more interesting, and more important the task.

The scientific rewards that accrue from the idea that the implementation of such strategies should make re-
searchers more careful and thoughtful not only in how appeals for volunteers are made, but in planning the research. For instance, if researchers tell participants as much as possible about the significance of the research, as though they were another granting agency—which in fact they are, granting researchers time instead of money—the emphasis is placed more heavily on the scientific rewards of doing important rather than trivial research.

Researchers who study clinical populations are bound generally to restrict participation to those who agree to participate. Rarely, an exception is made when the question to be studied is one that effectively precludes consent but is of such major public health concern that an institutional review board (IRB) will waive the requirement for consent. Salzman et al. (1991) compared two pharmacologic agents for management of disruptive psychotic behavior in a psychiatric intensive care ward. Patients whose behavior was self-destructive or dangerous to others routinely received one of these medications. The question posed was whether the drugs differed in terms of efficacy or side effects. The requirement of written informed consent would have precluded the research. A dual level of review by the local IRB and the state Department of Mental Health waived the requirement for prior informed consent. The study met two perceived criteria for waiver of individual informed consent: (a) the research question had high public health significance, and (b) the conduct of the research did not increase the level of risk to participants, because patients would have received one of the two medications studied. Studies such as that conducted by Salzman et al. are extremely rare. Virtually all research with clinical populations is conducted within the framework of individual informed consent to ensure the voluntary nature of participation. As is true in other research contexts, the voluntary nature of participation ranges from general announcements of the availability of participation to screening of all potential subjects followed by direct, and sometimes intense, efforts to encourage or urge participation by those who meet criteria.

In clinical research, the identification of a potential subject pool creates a sampling frame that allows comparison of the included volunteers with those who are not included in a study. Such comparisons are valuable to investigators, who are often able to compare systematically the included and excluded participants on a wide range of demographic or other characteristics.

There is one essential characteristic that distinguishes volunteers from nonvolunteers, and that is participation itself. Access to potential participants in clinical settings is often gained through the clinicians responsible for treatment. This poses a potential bias in recruiting. Many clinical researchers report informally, for example, that although patients are recruited from large numbers of wards, a disproportionate number of participants come from only one or two of them. Spohn and Fitzpatrick (1980) compared participating and nonparticipating subjects representing successive screening and self-selection in a population of schizophrenic patients being considered for a study of medication withdrawal. Patients who were research eligible were divided into those whose treating teams allowed them to be approached and those whose teams did not believe they should be approached. Furthermore, individuals asked to participate were divided into two groups—those who ultimately agreed and those who did not. Subsequent comparisons among groups revealed substantial differences between the original reference sample and the final consenting sample that, the researchers concluded, limit interpretation of data from the actual study.

In another example, Schooler (1980) identified a population of schizophrenic patients and compared participants in a clinical trial of injected versus oral medication with those who refused to participate because it was an experiment. There were no differences between the groups in relevant background characteristics. The hypothesis tested was that patients who had to take medication daily would relapse sooner than those who received medication by injection every few weeks, because the oral medication group would forget and stop taking their pills. The null hypothesis could not be rejected.

But, if the potential subjects in Schooler’s (1980) study who refused to participate in experiments tended to be those who are not compliant medication takers, the experiment could have been seriously flawed. Schooler tested this possibility by assuming that all who refused to participate were included in the study. If randomized to oral medication, they would have stopped taking it and relapsed. If randomized to injectable medication, they would not have relapsed. A comparison of relapse rates based on this hypothetical subject pool, not biased by exclusion of refusers, showed no significant differences between the groups. Thus, detailed descriptions of individuals who are excluded from clinical studies may provide valuable research rewards that increase knowledge and allow the test of focused hypotheses about potential effects in nonvolunteers. Such detailed comparisons are too rare in the clinical literature and are another example of investigations that warrant further study.

In summary, the psychology of recruiting participants for a research protocol is not dissimilar from other social marketing situations. There is a gray line between applying pressure to participate and being a competent recruiter and researcher. The gray area creates the opportunity for many ethical dilemmas. For example, is it ethical to employ young, attractive, verbal, and intelligent assistants of ethnic backgrounds similar to the target population in order to recruit participants? Participants may be trying to please the recruiter. Participants are also more likely to participate if they are familiar with the research assistants. Is it ethical to have research assistants spend time in the recruitment site, building a positive reputation, so that the potential participants are familiar with the researchers? Such a strategy is likely to increase the recruitment rate, but is that undue pressure? May a researcher ethically reward children with a classroom party for returning parental consent forms for a research proj-
ect, the reward contingent on return of the consent form whether or not the parent granted approval? Is it ethical for a researcher to convince a school district to adopt a proposed intervention as a school-wide curriculum? Evaluation of a school-wide program in collaboration with the district does not generally require parental consent. However, a researcher initiating and evaluating a similar program does require parental consent and extensive review by an IRB. The extent to which researchers must divorce themselves from the recruitment setting and restrain from creating an environment that enhances recruitment is unclear ethical territory.

Conclusion

This article is meant to raise more questions than it answers. As in the committee's previous article (Grisso et al., 1991), we have highlighted a number of issues because they reflect the special interests and backgrounds of those serving on the CSR, not because they are any more pressing than the many other questions about ethical standards facing research psychologists (cf. Pope & Vetter, 1992). The primary objective of this article is to continue the discussion and view of ethics as presenting opportunities for scientific rewards in psychological research with human participants.

In future articles, the CSR hopes to explore criticisms of review boards as overly zealous in exercising their gatekeeping function at the expense of scientists, who also have the ethical imperative to do sound research (e.g., Ceci, Peters, & Plotkin, 1985; Rosenthal & Rosnow, 1984). As Darley (1980) stated, there is an ethical imperative in doing sound research, for otherwise "we leave those who are attempting social change the prey of hucksters who are willing to put forth undocumented claims based on inadequate evidence" (p. 15). The study of this issue is meant to shed light on the difficulties currently faced by review boards (institutional or governmental) in assessing the rewards and potential conflicts of ethical choices in human subjects research.

REFERENCES


Harris, B. (1988). Key words: A history of debriefing in social psychology. In J. Morawski (Ed.), The rise of experimentation in American psychology (pp. 188–212). New Haven, CT: Yale University Press.


Macklin, R. (1990, September). Ethical dilemmas. Presentation to the


