The Institutional Review Board as a Mirror of Scientific and Ethical Standards

Ralph L. Rosnow, Mary Jane Rotheram-Borus, Stephen J. Ceci, Peter David Blanck, and Gerald P. Koocher

Decisions by institutional review boards (IRBs) are presumed to reflect the norms and standards of the scientific community. Such criteria have shifted as changes have occurred in experimental interventions and protocols, codes of federal regulatory agencies, norms among investigators, and expectations of participants. The tension created by shifting norms and standards raises two questions: (a) Should IRBs evaluate the scientific (e.g., design) features of the proposed research, and (b) should consistent standards be expected even in areas that are in constant flux (e.g., AIDS research)? We discuss these questions and propose a mechanism to keep IRBs abreast of emergent issues and sensitized not only to the costs of doing research but also to the costs of not doing it.

Over the past two decades, psychologists doing research in a number of areas have witnessed a proliferation in the ethical standards to which they are held accountable, especially when the research involves human participants. Because of changes in ethical standards commensurate with changes in scientific practices, psychologists not only must be concerned with the protection of autonomy, privacy, and justice in the process of recruiting research participants but must ensure that the highest standards of ethical and scientific conduct will be followed throughout the research process (Christakis, 1988; Williams, 1984). Given the broad range of scientific and ethical challenges punctuating the social and scientific zeitgeist, the Board of Scientific Affairs (BSA) of the American Psychological Association (APA) revised the name and functions of the former Committee for the Protection of Human Participants in Research (CPHRPR) to become the Committee on Standards in Research (CSR; see Grisso et al., 1991).

This change in title and purpose was intended to expand the area of concern in which the CPHPR had previously operated. It would now include (a) promoting freedom of research consistent with the highest ethical and scientific standards, (b) monitoring attitudes and concerns regarding the use of human and animal participants in research, (c) preparing written and oral statements relevant to research ethics and scientific conduct, and (d) updating the Ethical Principles for the Conduct of Research with Human Participants (APA, Committee for the Protection of Human Participants in Research, 1982). Recently, the CSR invited suggestions and experiences from psychological researchers and practitioners, editors of APA journals, and other interested parties to prepare for the task of revising the APA's document on ethical principles for research using humans (Mitchell-Meadows, 1992). The CSR has also sought to foster the view of research ethics as an opportunity for increased scientific and societal rewards rather than as an affront to the integrity of sound research (Blanck, Bellack, Rotheram-Borus, & Schoeller, 1992).

In this article, the CSR continues its public discussion of problems that it believes are in need of study. The general issue addressed here concerns criticisms directed against institutional review boards (IRBs) as overly zealous in exercising their gatekeeping function at the expense of scientists, who also have the ethical imperative to do sound research (see Ceci, Peters, & Plotkin, 1985; Rosenthal & Rosnow, 1984). By opening this issue to discussion, we seek to accomplish three goals:

First, we hope to shed light on the difficulties faced by IRBs in assessing the rewards and potential conflicts of ethical choices in research with human participants. The title of this article encapsulates the idea that IRBs are presumed to mirror the incipient ethical dilemmas facing the scientific community and, in the process of evaluating risks and benefits, to mirror the current standards of scientific practice. As the number and sheer
complexity of dilemmas increase, the role assigned to IRBs could become more multifaceted.

Second, we begin to explore how changes and emerging trends in the ways that IRBs choose to perform their role may affect freedom of inquiry in science. Just as standards of research practice have undergone evolution, so have the ethical criteria implemented by IRBs. To illustrate this point, we examine some of the difficulties faced in establishing consistent ethical procedures for research on the acquired immunodeficiency syndrome (AIDS) pandemic. Researchers in other areas of psychology have also experienced paradigmatic changes, but perhaps no area of human subject research has been more affected by change than AIDS research.

Third, we aim to stimulate interest in the development of mechanisms to guide and monitor the IRB process. In so doing, we strive to ensure that there will be no cessation of studies that need to be done in order to answer important scientific and societal questions. We believe there are potential gains to be realized by both science and the general public as IRBs become more sensitive to the idea that the failure to conduct a study that has been proposed is as much an act to be evaluated on ethical grounds as is the conducting of a study (Rosenthal & Rosnow, 1984; Rosnow, 1990). We also hope to facilitate more consistent implementation of ethical standards across a field of research.

Assessing Risks and Benefits in Human Research

A central responsibility of IRBs is to ensure that the potential benefits to the individual research participants (and to society) will be greater than any risks that may be encountered by participation in the research (Stanley & Sieber, 1992). Before the 1950s, researchers were not held to a systematic evaluation of whether their studies met this requirement. There were no IRBs and few standard practices against which to assess whether the investigators were operating in an ethical manner. A radical shift occurred in 1954, when the United States Public Health Service decided to organize a large-scale experiment to test the efficacy of the Salk vaccine to protect children against polio.

Involving almost two million children, this was the first large-scale field trial ever mounted, and it raised a number of controversial issues (Meier, 1989). A basic concern was whether to use a randomized design that would place untreated (control) participants at greater risk than treated participants or to give all the children the poliomyelitis vaccine. It was decided that the only scientifically acceptable way to find out whether the Salk intervention can be detected. It is not difficult to imagine how the strict application of standards developed in drug trials could pose problems for researchers in other areas. On the other hand, not including design issues as a prerequisite for approval could have negative ramifications (Meier, 1992). For example, if a large clinical trial were proposed to examine the efficacy of an experimental drug for halting progression of fatal illness and prolonging a high quality of life, the design must allow clear conclusions to be drawn. The failure to meet standards of control and sample size in such trials could result in a waste of governmental, institutional, public, and personal resources.

To be sure, the criteria originally invoked by IRBs continue to evolve, in part because of pressures by advocacy groups. For example, groups have lobbied for quicker release of relevant drugs for dying patients, drug companies have marshaled support for changes in access of research participants to experimental treatments, and researchers who see deficiencies in the medical community's and general public's understanding of the meaning of informed consent have pressed for education of these constituencies (Mitchell & Steingrub, 1988). Such pressures can produce a ripple effect leading to changes in national and regional criteria for the evaluation of research, ultimately affecting IRBs that are obliged to incorporate such standards.

The criteria used by medical school IRBs to address the issues with which they grapple in evaluating clinical trials have also come to be applied to behavioral and social science (Cann & Rothman, 1984). With the introduction of these criteria into a wider sphere have come concerns by psychologists as to the potential for misapplications (Ceci et al., 1985). One issue is whether evaluating the technical (not just the ethical) merit of a study is within the purview of IRBs. Historically, the role of review boards was to ensure that informed consent was obtained, the confidentiality of participants was safeguarded, and the recruitment process allowed equal access to all relevant potential participants. The objective was to preserve the autonomy of potential participants to the maximum extent possible. In recent years, the role of IRBs has often been expanded to include issues not specifically related to the autonomy of research participants.

A broad range of design issues is now included in many IRB discussions. The design of a drug trial is generally viewed as deficient unless there are adequate controls, random assignment, and a sufficient number of observations to ensure that the results due to the experimental intervention can be detected. It is not difficult to imagine how the strict application of standards developed in drug trials could pose problems for researchers in other areas. On the other hand, not including design issues as a prerequisite for approval could have negative ramifications (Meier, 1992). For example, if a large clinical trial were proposed to examine the efficacy of an experimental drug for halting progression of fatal illness and prolonging a high quality of life, the design must allow clear conclusions to be drawn. The failure to meet standards of control and sample size in such trials could result in a waste of governmental, institutional, public, and personal resources.

One question we raise is whether the composition of IRBs should be systematically expanded to include persons with expertise in areas not specifically directed to the autonomy issue. Should the membership include people who are sensitive to nuances in the interface of ethical and technical aspects of behavioral and social research (see, e.g., Suls & Rosnow, 1981)? Even if psychological researchers were to support this idea, considerable variability would exist among IRBs in deciding whether issues of design and methodology fall within their re-
sponsibility. Among the members of the CSR, there is also disagreement over whether technical aspects of the research are a proper concern of IRBs or whether such scientific issues should be resolved before submission to the IRB. Suppose an IRB insisted that a psychological study be radically redesigned to conform to criteria that the researcher, with a track record of successes in the area, viewed as an unrealistic demand. What recourse does the researcher have to press his or her claim or to resolve the disagreement expediently?

**Maintaining Ethical Standards in a Changing Field**

In practice, IRBs in medical schools are likely to be more intrusive in their critique of scientific merit than are review boards located in academic departments in liberal arts colleges. This greater intrusiveness is predicated on the belief that there is generally greater potential risk to participants in clinical trials and studies using invasive therapeutic procedures than in those using traditional behavioral procedures. Since the identification of AIDS in the early 1980s, the virus has presented a series of challenges to researchers attempting to maintain integrity in the application of ethical and scientific standards (Bayer, Levine, & Murray, 1984). The social stigma of AIDS, as much as the association of AIDS and death, intensifies the challenge (Bayer & Gostin, 1990). In the context of a life-threatening disease, it is critical that the principles of privacy, justice, and autonomy be consistently applied at every stage of the research. However, the definition of ethical practice and researchers' ability to ensure confidentiality and justice among human immunodeficiency virus (HIV) infected persons has changed over the last decade.

To illustrate, when an IRB makes decisions about maintaining confidentiality regarding a communicable disease that poses a substantial threat to the public's health, there is inherently a tension between an individual's desire to control personal information and society's desire to gain access to information (Bayer et al., 1984). The strain is increased because those at risk are often politically, socially, and economically stigmatized. In the midst of the AIDS pandemic, it is often not clear that researchers can ensure the confidentiality of their participants in research. Even researchers who have obtained "certificates of confidentiality" from the National Institutes of Health (NIH) ponder the consequences should the issue of confidentiality be challenged in the courts. Confidentiality is not the only source of tension, nor are issues of confidentiality confined to HIV research (Blanc et al., 1992), but the tension seems almost palpable in the case of HIV research.

To give another example, the availability of prophylactic treatments has led to increased advocacy for HIV testing, especially of persons considered at high risk (Bayer & Toomey, 1992; Kutchinsky, 1988). Behavioral researchers working in this area are increasingly interested in evaluating the consequences and reactions to testing. However, their ability to protect the confidentiality of those who are tested has decreased. Knowing one's serostatus is seen as the basic first step to adopting safer sexual practices and, therefore, a desirable goal of public health officials. There is conflicting evidence that at least raises the possibility that those who know they are HIV seropositive are not necessarily more likely to adopt safer sexual practices or to disclose to sexual partners (Fox, Odaka, Brookmeyer, & Polk, 1987; Goedert, 1987; Ostrow et al., 1989). Given that it is seen as good public health policy to know one's results and that some states have mandated the reporting of serostatus, can research studies of HIV testing be conducted without jeopardizing the person's right to privacy and confidentiality concerning serostatus? Analogous issues have been raised in other areas of research in psychology, such as the problems in protecting confidentiality in field studies of child abuse or venereal disease (Blanck et al., 1992).

Indeed, the question of HIV testing becomes even more complex when children considered at high risk (e.g., runaway, gay-identified) are the focus of the study (Rothram-Borus & Koopman, 1992). In some states, parents can mandate that their children be tested or that they as parents be informed of the results of testing. If adolescents participated in the research as mature or emancipated minors, they nevertheless may return home at some point in the future and be under the supervision of their parents. Are parents then entitled access to information obtained by researchers when the adolescent held a different status? New York State Public Health Law (New York State Laws, 1988) mandates that serostatus be recorded in medical charts. This law also stipulates that physicians be allowed to disclose serostatus to parents when necessary for care or treatment unless the youth already has the legal authority to consent to health care (see Greater New York Hospital Association, 1988). Even if adolescents remain out of their homes, recording information about serostatus in a medical chart opens the opportunity for disclosure to parents. Studies evaluating the impact of HIV testing could create nightmares for investigators who strive to maintain consistent ethical practice.

A different type of ethical dilemma was created for IRBs when HIV serological testing emerged in the 1980s. Homosexual research participants were enrolled in ongoing protocols when the ELISA (enzyme-linked immunosorbent assay) and Western blot test for HIV became available (Holder, 1985). Researchers in several studies assessed the serostatus of their participants in the ongoing longitudinal studies with informed voluntary consent. The researchers offered participants the option of knowing their serostatus, but participants could choose not to be informed of the results of their tests. Some participants agreed to be tested but did not want to know their status. Within several years, the NIH adopted the policy that persons tested must be informed of their serostatus in the interests of the general public's health.

Consider the implication of such a ruling for participants who agreed to be tested under one set of guidelines if the guidelines later shifted. Given such a ruling, how was the researcher to handle this situation when partic-
participants had agreed to be tested, were tested, and did not
return for the results of their serostatus? Was the re-
searcher required to inform the participants? Should the
conductor conduct studies that evaluate the impact of HIV testing
even though this testing is not linked to ongoing health
care? Are there hypotheses in social psychology, clinical
psychology, developmental psychology, and so forth that
cannot be addressed because of the researcher's inability
to protect confidentiality? Questions such as these reveal
how difficult it can be to maintain a sure moral footing
regarding issues of autonomy and confidentiality when
there is a slippery ethical slope.

Inconsistencies in Decision Making

There appears to be great variability in the standards in-
voked, and in turn the recommendations put forward,
among IRBs (Ceci et al., 1985; Prentice & Antonson,
1987). It has been reported, for instance, that although
IRBs are relatively successful in ensuring privacy and
overseeing consent, they are less effective in weighing risks
and benefits (Williams, 1984). Inconsistencies in evalu-
ating risks and benefits may stem from biases in the as-
se ssment of protocols, the composition of review boards,
and the nature of committee action and interaction. The
point is that different standards are being applied to re-
search at different institutions and in different parts of
the country. Inconsistent standards create the appearance,
if not the possibility, of injustice (Rosenthal & Blanck,
1993). In one case, it often happened that the identical
proposal that was approved without amendments at one
university was amended or even disapproved at a nearby
university in the same city (Ceci et al., 1985).

Kimmel (1991) asked a sample of psychologists to
give their opinions about the ethical costs and benefits of
hypothesised studies. He reported that respondents who
tended to put a greater emphasis on research costs were
primarily female; had recently received the PhD degree
in counseling, school, or community psychology; and were
employed in service-oriented contexts. By contrast, re-

spondents who tended to put a greater emphasis on re-
search benefits were primarily male, had held their doc-
torates in a basic psychology area (such as social, exper-
imental, or developmental psychology) for a longer time,
and were employed in research-oriented contexts. These
results raise the suspicion that individual biases toward
costs and benefits due to the composition of an IRB could
influence how it ultimately decides particular cases (see
Ceci et al., 1985; Hamsher & Reznikoff, 1967; Kallgren

State laws that limit the types of information and
degree of acceptable risk to research participants are an
indirect source of variability among IRBs (e.g., New York
State 14 NYCRR 527.10). To the extent that a state insists
on such limitations, IRBs will be obliged to impose stricter
standards, and, in turn, there will be fewer opportunities
to conduct research of a critical nature in that state. In
the case of AIDS research, such regional variability is
likely to influence the selection of sites for vaccine trials.

At present, there are several potential drugs that may
be tested in AIDS vaccine trials (Taylor, 1992). To be
useful, the trials must be conducted with populations that
have a current conversion rate of seropositivity of about
3% a year. This implies that the trials be conducted on
only those individuals who are at highest risk, which in-
cludes those who share intravenous drug needles and
young men who identify themselves as gay. Large numbers
of participants will need to be recruited in each locale.
In addition, it will be necessary to ensure that participants
recognize that receiving an experimental vaccine does
not imply that they are protected against HIV. Also, there
are unlikely to be any biological markers to signal whether
any individual is immune to the virus.

These are rigorous constraints under which to mount
a vaccine trial. Although community advisory boards en-
sure that the IRBs mirror local standards and norms
(Valdiserri, Tama, & Ho, 1988), wariness and mistrust
of scientists—particularly among minority participants
at high risk—could be substantial in view of the experi-
ence in previous circumstances (Ad Hoc Advisory Board,
1973; Thomas & Quinn, 1991). Confronted by these
constraints, it is not surprising that researchers find it
tempting to carry out such studies in regions where they
need not worry about community advisory boards, such
as in the developing world rather than in industrial coun-
tries. However, think of the untold new ethical dilemmas
if developing countries should become a testing ground
for the rest of the world.

Investigators distressed by the actions of local IRBs
often voice the complaint that the committees act as a
police force rather than as a protector of the rights of the
participants (Christakis, 1988). Those researchers con-
ducting low-risk interventions or epidemiological surveys
have felt especially burdened by the demands of IRB pro-
cesses (Cann & Rothman, 1984). Protocols of low risk
are actually excluded from the necessity of undergoing a
full IRB review if they do not violate one of three basic
criteria: (a) anonymity of responses, (b) absence of civil
or criminal liability, (c) sensitive aspects of behavior (De-
partment of Health and Human Services, 1981). Most
epidemiological surveys are exempt from review by IRBs,
given these criteria. However, concerns regarding liability
and shifting standards for the protection of human par-
ticipants can mean that the same criteria are being applied
to noninvasive interview surveys as to invasive medical
procedures (Cann & Rothman, 1984). Thus a further
challenge is to prevent the erosion of confidence among
researchers, particularly those engaged in low-risk studies,
who may perceive the review process as arbitrary or even
irrational.

Some of the difficulty stems from local practices that
are inconsistent across institutions, and some of it derives
from conceptual confusion at the national or policy level.
Regarding the latter, consider the case of educational re-
search. Under some circumstances, research can be recast
as curricular or educational and may thereby qualify for
either expedited review or no review at all. School district employees frequently implement new curricula or learning technologies, sometimes replete with a formal evaluation plan, and are not bound by the same standards that apply to university researchers who propose similar curricula or technologies. Major large-scale educational interventions that might produce a negative impact on the lives of children are implemented with little difficulty. Witness the introduction of the “good touching–bad touching” curriculum, which, its proponents claimed, would increase the reporting of sexual abuse, although others have argued otherwise (see Reppucci & Havgaard, 1989). Another example was the new-math curriculum, which is now widely viewed as having set back an entire cohort of youngsters.

More recent illustrations include drug-related skills (e.g., Hawkins, Catalano, & Kent, 1991) and sexual-health curricula (e.g., Patierno & Britton, 1992) aimed at high-risk behaviors. Notwithstanding the absence of pilot data to justify the implementation of educational interventions such as these, proponents are rarely held to a level of cost–benefit analysis comparable to the most innocuous list-learning study proposed by an MA student in experimental psychology. Put baldly, there is a perception of hypocrisy, in that researchers are being asked to justify relatively innocent procedures while others are allowed to pursue potentially damaging practices with little or no justification. If protecting children from risk of harm is a concern for university researchers, then fairness would seem to dictate that it be a concern of school employees as well. Similar examples can be found in common medical practices, the behavior of attorneys toward clients, and so forth. Researchers feel singled out.

In addition to a discussion of community standards in its deliberations, IRBs should be acquainted with national disciplinary standards. At present, such standards are not clearly articulated, but they can be surmised from an analysis of published articles. APA journals are replete with studies that entail deception, psychological stress, and deliberate avoidance of debriefing (e.g., where researchers fear that debriefing could lead to unwanted social repercussions, such as a diminution of helping behavior among subjects in a bystander apathy experiment). There needs to be a balance between local standards of IRBs and national standards because investigators have a foot in each of these communities.

**Toward Cost–Benefit Identifications**

In view of the shifting standards governing research and the unevenness with which different IRBs function (Ceci et al., 1985), a process is needed to maximize consistency across IRBs and at the same time keep them abreast of emergent issues. For a variety of reasons—not the least being the need to couch deliberations in the context of local community standards of ethics—no top-down national regulatory body is desirable. Instead, it would be valuable for IRB members to be supplied with a casebook of actual research protocols that have received extensive review and analysis by social scientists, bioethicists, and research participants (both investigators and subjects). Such a casebook should aid IRB members in making explicit the costs of doing research as well as the costs of not doing it. It might also give them an opportunity to expose their own ethical biases to scrutiny so that they can function judiciously and equitably.

Although a casebook ought to help sensitize IRB members to relevant costs and benefits, there will always be cases that require extrapolation and novel analysis. This raises the possibility of conflicts among the opinions and values of IRB members, investigators, and institutions. Although this is unavoidable, it could be alleviated by the creation of an advisory board within the Office for the Protection of Subjects From Research Risk. Parties to disagreements could request an analysis and review of an IRB decision by such a board. Its analysis could serve merely an educative function or, in cases in which all parties gave previous stipulation of their willingness to accept the analysis as determinative, it could be binding.

A description of the mechanics of such a board (i.e., its creation, authority, membership, and operating procedures) is beyond the scope of this article, but we raise the idea here for further discussion.

**Conclusion**

We have looked at how the role of IRBs has expanded over the past two decades and the ways in which inconsistencies seem to adhere to decision criteria that mirror changing standards of fields that are in constant flux. Clearly, it is critical to establish guidelines for the evaluation of research protocols (Prentice & Antonson, 1987). There must also be a way to limit the power of review boards (instead of adopting the strategy of an ever-increasing role), but current proposals include evaluation of the technical merit of the study as one component of evaluating the risk–benefit ratio for involvement of human participants. Finally, we raised the idea of providing IRBs with a book of case studies and accompanying ethical analyses to sensitize members to troublesome issues and nuances in the behavioral and social sciences. Central to such case studies would be an explicit enjoinder that IRBs take into consideration not only the costs of doing the proposed research but also the costs of not doing it. In general, we believe that a mirror must also be held up to the review process, so that its existence, fairness, and effectiveness can be examined and justified (see Cowan & Adams, 1979; Hershey, 1985).

**REFERENCES**


