

ETHICAL PROBLEMS IN EVALUATION RESEARCH*

Elisabeth J. Johnson

For some time, social scientists have been concerned with ethics in research. Much of what they have written has focused on research in general, but many of the points they raise are applicable to evaluation research in particular. Included among these concerns are informed consent, the right to treatment, the right to refuse treatment, implicit coercion, powerlessness of subjects, and, perhaps most important, external access to confidential data.

Evaluation studies approach other ethical domains as well, such as responsibility to funding sources, the threat to jobs of persons working on programs being studied, and the needs of the community as perceived by its residents. Many of the issues are inextricably tied to political concerns; we are concerned here with politics at the agency and specific project levels. This presentation shows that these issues are so complex and so important to the outcome of the research that strategies for dealing with them should be developed in the initial planning stages of the study, along with other methodological procedures.

Position of Relative Power of Researchers and Subjects¹

Researchers should begin by considering their positions as they relate to research participants. The researcher has a number of advantages. First, the person comes with the sanction of whatever institution is sponsoring the study by virtue of having had the proposal accepted. This sanction is not an inconsiderable power source, for such funding agencies as part of "the establishment" represents high authority, a fact that is not lost on program participants. Second, the investigator has the advantage of superior knowledge in being the only one who knows the grand design. The evaluator is, therefore, in a position to manipulate variables, alter the design, and give the orders. The researcher continues to gain superior knowledge through education and training not only in the substance of the specific discipline but also in functioning with "establishment" institutions.

* Accepted for publication April 26, 1980.

Minority researchers share some of the same power advantages that their majority colleagues enjoy. And investigators who come from minority groups will likely have "the establishment" perspective, because of their training. This is not to say that their value as social scientists should be minimized; quite the contrary, for their membership in minority groups should give minorities an understanding that many middle class whites would find difficult to achieve.

Research subjects are in quite a different power position. Evaluation research is usually focused on action programs designed to ameliorate the consequences of social injustice. Program participants typically come from groups which have been the victims of injustice. These victims include racial minorities, welfare recipients, the elderly, children, prisoners, ex-convicts, and mental patients, all of whom are powerless in relation to the larger society. Participants in action programs are usually beholden in some way to "the establishment" for the services they receive, and there is an implicit threat to these services if participants refuse to cooperate with evaluation efforts. Such persons do not usually challenge the established order and there is no reason to expect that they would do so in response to a research program.

Privacy and Confidentiality

Herbert C. Kelman has pointed out that one of the principal aspects of human dignity is the right of choice. The choice to reveal or keep oneself private is cherished by all persons, regardless of social class. It is interesting to note that privacy as a phenomenon increases with class privilege. For example, executive bathrooms are locked; being awarded a key is a sign of increase in status, but the workers' bathrooms feature multiple stalls with only the most rudimentary concealments. This differential is also true in terms of research, with upperclass people reserving more of the right to withhold information about themselves than the lower classes are able to do.² Persons who cannot protect their privacy legitimately will do so deviously; this behavior could destroy a research project.

One of the major ethical issues confronting researchers is confidentiality and privacy. The fact that subjects may feel powerless in the face of research plans in no way implies that they do not feel exploited and exposed. What is worse is the fact that subjects have no choice in whether the program will be evaluated or not. The implications for the study's success are serious, because the participants' lack of influence over the program makes the evaluation vulnerable to indirect manipulation through clients' lying and other

forms of sabotage.

Revealing certain types of information can pose a definite threat to certain populations such as prisoners and welfare recipients. Promises of confidentiality, meant sincerely, are usually given by evaluating groups. However, it is now impossible to give that promise with complete assurance that it can be kept, given computer technology and data banks which make information retrieval very simple for knowledgeable persons. Hoffman and Miller have described, in rather alarming terms, a method of getting private information on a given individual from a data bank using readily available information such as occupation and city of residence.³

The federal government maintains data banks which contain information about clients likely to participate in action programs. An example is the CODAP system developed by the National Institute on Drug Abuse (NIDA). NIDA requires all drug programs which receive federal funds to forward certain information about their clients for inclusion in the CODAP system. These clients are truly at risk, for many of them have been in trouble with the law. The conflict arises when there must be an accounting of the expenditure of tax dollars and evaluation researchers are the ones who do it.

The issue of maintaining data confidentiality has led to some troublesome times for some programs. In the Chicago Woodlawn Project, which worked with the Blackstone Rangers and the Devil's Disciples, notorious youth gangs, an evaluation was terminated over the issue of confidentiality. The program was considered by some to be novel and successful, because it worked through the structure of the gangs to achieve change in the community. During the violence that followed the assassination of Martin Luther King, Jr., there was significantly less violence in the Woodlawn district than in other Chicago neighborhoods. Yet the program was controversial in many ways: Mayor Daley opposed it because he could not control the funds and a senate sub-committee wanted an accounting because the program worked with criminals. The evaluators were having difficulties in getting the staff and clients to cooperate with the study because of all the outside pressure. Ultimately, the senate sub-committee subpoenaed the raw data which had been collected at a time when the program was troubled. When the evaluators delayed releasing these data, the sub-committee ordered a site visit. During the visit, some documents from the program were taken by a committee staff member, presumably inadvertently, and later returned with a note saying that they had been copied.⁴ Incidents of this nature are not infrequent, especially in controversial programs. Usually the researchers get blamed for this violation of confidentiality

even when it is clearly not their fault. Program evaluators must find ways to avoid these pitfalls in order to gain and maintain the confidence of program staff and participants. One way to do this is to plan and carry out research in partnership with consumer representatives.

Political Interests

Action programs tend to be favored by neighborhood groups, human service workers, and social scientists. There are, however, many groups with competing interest and perspectives. Money is often given directly to an agency or program, by-passing state and local authorities. Government officials are, at the very least, skeptical of such an arrangement. Legislators who appropriate money through Health and Human Services, the Justice Department, or the Department of Labor are obliged to have an accounting of such expenditures to their constituents. Ordinary middle class citizens feel resentful about efforts to pull minority citizens into the mainstream, for every newcomer who enters the job market is seen as a potential threat to their interests. These conflicting interests cause incredible tensions within and around community programs.

The demand for cost effectiveness is not really understood at the program level. The funding source wants to demonstrate efficacy as soon as possible so that they can justify further spending. It is nearly impossible to work under such pressure: trying to produce results in too short a period of time, with difficult clients, and, often, inexperienced staff. Staff dissension abounds, clients drop out of the program in frustration, funding is threatened and it is in this context that the evaluators are expected to commence their studies. It is understandable that program staff feel beleaguered and defensive and do not willingly cooperate in a program's assessment. Under these conditions, evaluation cannot possibly measure the program at its potential best, especially because, as C. Weiss points out, such studies have a tendency to find negative results.⁵

Weiss suggests that timing is crucial for meaningful evaluations. Many studies are started before a program has worked through its problems and adequately trained its staff. Premature evaluations often lead to abandonment of a program before it has had a chance to prove itself. Another problem that funding institutions needs to consider is that no one program can remediate all ills in society. The best we can expect is the amelioration of the social problem for a small group of victims. Consequently, large impact studies are probably inappropriate.

Consequences of Negative Evaluations

In the final analysis, a negative result is the consequence most feared from evaluation. When a new program is funded, it sets up a chain of events that have long-reaching effects on a community. A needed service is provided and local residents come to expect the service to be available indefinitely. Jobs are created and they are often filled by neighborhood folk who might never have obtained similar positions elsewhere. The program, if it meets certain criteria, can become a training site for graduate students, thereby increasing the quality and quantity of the services offered. A negative evaluation can destroy such a network. Evaluations that focus more on the process and milestones of the project might be seen as more helpful and useful in decision-making and change. The agency would then have a better chance of defining itself and achieving its goals.

Many social action programs were started as alternatives to traditional services. This suggests that new methods are being used, most of which have not been tested. Some failures are unavoidable as the staff tries out new ideas and methods in their search for the most effective models. A negative evaluation stifles this creativity; in response, the staff resorts to the same old tried methods which never really worked, but offer a measure of safety, nevertheless. *Creativity is risky and evaluation ought to accomodate to that fact.* Self-help originality can blaze a trail where "establishment" programs have failed precisely because the workers will not have the same biases as the professionals and academicians. To evaluate such programs, a new approach and perspective is needed.

The Right to Treatment, The Right to Refuse Treatment, and The Loss of Temporary Benefits

Many evaluation studies use the experimental method or aspects of it, i.e., they use control groups, give and withhold treatment, offer temporary benefits, and then study the effects of these manipulations. An example is the Negative Income Tax Experiment in New Jersey which gave income supplements to persons at the lowest end of the tax brackets. The grants were given for a specified time following which recipients were evaluated to measure the many benefits derived from the program and then compared to a similar group who did not receive the extra money.⁶ It can be argued that persons who receive this benefit become accustomed to a new lifestyle and therefore the research causes the subjects to feel worse after completion of the study than they did before. One could also argue that to use a control group of needy persons and exclude them

from the cash benefits is inherently exploitative. There are no neat solutions to these dilemmas, but researchers would be remiss, indeed, to rush into such a project and neglect to consider the total impact of such activities on the population.

In other cases, special treatments may be given to patients in health clinics. The staff has reason to believe that the treatment will be effective and is, therefore, reluctant to exclude some patients in order to establish a control group. If the benefit to the patients from the experimental treatment will be major, the researcher should seek an alternative design. This situation is one of the major arguments for doing quasi-experiments in clinical settings; quasi-experiments eliminate the necessity of randomization and use of control groups.⁷

Informed Consent

Most research projects require that a subject or a guardian must sign an agreement to participate in the program. Such agreements must contain enough information about the study so that all risks are made clear. The idea is to help subjects make intelligent decisions about participation or not. This process is called informed consent. In basic research experiments, this issue is obvious and the requirements, while possibly complicating the design, are quite clear. However, in evaluation research, the need for informed consent and the form it should take is ambiguous at best. There are numerous points to consider.

In the best of circumstances, the research is compromised if the subject knows too much. The outcome of the treatment can easily be attributed to the demands of the situation. This is especially true in action programs where the subjects have as much stake in looking good as do the staff. Another problem is the Hawthorne effect; just participating in a study can produce effects on individuals which obscure the treatment effects.

Another aspect of informed consent is related to the power issue. Populations in action programs are receiving benefits which may make them feel beholden to the service provider. In such a situation, the individual is subject to implicit coercion. Although information about treatment effects, the reason for the study, and the voluntary nature of participation are explained to the subject, there remains unavoidable pressure to comply. In cases like this, the informed consent agreement meets the requirement of being voluntary, but violates the spirit of the law to a degree. This is especially true in programs involving prison populations where the coercion is not so

subtle and the benefits of cooperating may be as obvious as reduced sentence time.

The final point is also related to the power issue. Even though the treatment and the study are explained to subjects, in most instances they are not able to understand it as well as the investigators or the program staff, either because they do not know as much about the topic or because they are not as well educated. The onus is on the researcher to make such information as can be shared as understandable as possible.

Some have argued that social experimentation, in contrast to biomedical research, does not pose hazards against which participants must be warned.⁶ However, in some cases subjects may be worse off if the treatment is ineffective; subjects might also develop fears and misgivings, needing reassurances from staff. The point here is that investigators should be alert and sensitive to these issues and also flexible enough to alter programs to accommodate those who perceive themselves to be at risk.

What Can Be Done?

It is clear that ethical issues in evaluation research are at best ambiguous, and at worse, serious enough to halt the research. There are some strategies which can minimize the risk that these problems present. At least two strategies are required for all research funded by Health and Human Services. All applicants for research money must describe the risk to human subjects and the steps taken to protect them from these risks. And each program is expected to organize a peer review group to analyze each potential hazard and to make recommendations for effectively reducing such problems as they identify. Many programs go further and submit their proposals to independent review groups.

These policies are excellent ways to reduce the risks to research participants, but the methods only satisfy a portion of the ethical problem. The notion of review groups could easily be expanded to include persons who are representative of the population to be studied. Kelman refers to "participatory research" as a way of including such individuals.⁹ When a researcher is designing a study, the researcher can include program staff and neighborhood representatives in the planning. In so doing, each group has a chance to raise questions that are of particular saliency. It also allows client advocates to articulate the fears and cultural concerns that the clients may experience. If researchers sincerely value this type of

cooperation, they may be able to capitalize on the ideas that **are** generated from staff and lay groups. Program staff have ideas **about** what they do that is effective and what the goals of the program should be. In participatory research, the investigators help the staff to structure the treatment protocol in such a way as to obtain the **best** result from an evaluation. Researchers must be honest **about what** research can and cannot do. Together the three groups frame the research questions. Since all have participated in the planning, all have an interest in seeing the study completed. An added benefit could well be that new ways of looking at old questions may emerge from the cross fertilization of differing perspectives. Participatory research can affect all the issues of ethics that have been raised in this discussion.

The issue of confidentiality and privacy is very complex. Some of the time, the strategies that have been mentioned will prove to be effective and guard against unwarranted intrusions of privacy. However, it is important to realize that review committees are subject to abuse and must be formed with the same care that is used in the research design. These committees are usually made up of other professionals (peers) and citizens of status from the community. It is argued that such persons are needed to provide the necessary expertise to make the judgments asked of them, which may be true in some cases. In the interest of fairness, persons who represent the concerns of the subjects and ordinary community folk need to be placed on these committees.

In a recent review process, one highly placed professional was discovered to be a member of three groups reviewing a research proposal, the clients of which were drug addicts in a treatment program. The proposal was approved; the professional in question is a close colleague of the principal investigator; there were no ex-addicts, community representatives, or treatment staff on any of the three committees. An objective review? Hardly! Yet, on paper, a three level review appears to be very conscientious and rigorous. Given the various risks drug addicts face regarding confidentiality, privacy, freedom of choice, and coercion, it seems imperative that advocates for their well-being be included in the review.

Other protections of privacy are more technical, such as setting a procedure for removing client identifiers from all information to be put in data banks. Indeed, only such information as is needed about a client to provide safe treatment should be obtained at all.

A final suggestion is general and long-term in nature: Social scientists have an obligation to educate the public to the nature of evaluation research and its danger to research participants. Although the state of the art is still at a primitive level, there is no reason why public groups and legislators cannot be informed about what is known. It is only in this way that we can begin to get funding institutions to accept and value the rights of clients regarding ethical issues.

Notes

¹Herbert C. Kelman. "The Rights of Subjects in Social Research in Terms of Relative Power and Legitimacy." *American Psychologist*. Vol. 27, No. 11 (November, 1972) 989-1016.

² Herbert C. Kelman. *A Time to Speak: On Human Values and Social Research*. (San Francisco: Jossey-Bass, Inc. 1968). See also, Barry Schwartz: "The Social Psychology of Privacy." *American Journal of Sociology*. Vol. 73, No. 6 (May, 1968) 741-752.

³Lance J. Hoffman and W. F. Miller. "Getting A Personal Dossier From a Statistical Data Bank." *Datamation*. (May, 1970) 74-75.

⁴John Walsh. "Antipoverty [Research and Development]: Chicago Debacle Suggests Pitfalls Facing OEO." *Science*. Vol. 165, No. 3899 (19 September 1969) 1243-1245.

⁵ C. Weiss. "Politicization of Evaluation Research." *Evaluating Action Programs: Readings in Social Action and Education*. C. Weiss, ed. (Boston: Allyn & Bacon, 1977).

⁶H.W. Reicken and R.F. Boruch. *Social Experimentation: A Method for Planning and Evaluating Social Intervention*. (New York: Academic Press, 1974).

⁷ T.D. Cook and D D.T. Campbell. *The Design and Conduct of Quasi-Experiments and True Experiments in Field Settings*. (Chicago: Rand-McNally, in press).

⁸ Reicken and Boruch. *Op. cit.*

⁹ Kelman *Op. cit.* (1972) 1012-1016.