

FACTORS IN THE PRACTICE, ORGANIZATION, AND THEORY OF EVALUATION

H. GIL PEACH

Pacific Power & Light

ERIC HIRST

Oak Ridge National Laboratory

ABSTRACT

This paper calls attention to the importance of a technical focus on organizational factors in evaluation design. Pro-active planning to influence the organizational structure of evaluation as well as the context and research climate of an evaluation effort, and the inclusion of multiple organizations with diverse perspectives can help improve the validity and usefulness of evaluation results. This paper discusses the aspect of the theory of evaluation and its application, based on our experience managing the evaluation of the Hood River Conservation Project.

THE "THIRD EXPERIMENT" IN HOOD RIVER

The Hood River Conservation Project (Project) was a test of the reasonable upper limits of a residential retrofit program, given current knowledge of residential weatherization materials (Hirst, 1987). From the beginning, the Project was multi-organizational and was premised on the mutual cooperation of the primary energy-planning organizations in the Pacific Northwest. The Project was unique in its focus on use of research methods to accurately evaluate the process and effects of a conservation program, its ongoing attention to these issues (manifested in the devotion of substantial resources to data collection and management), and its efforts to develop a consensus process among groups that were traditionally adversaries.

The Project was proposed by the Natural Resources Defense Council, funded by the Bonneville Power Administration, and operated by Pacific Power & Light (Pacific) in Hood River, Oregon with the cooperation and support of the Hood River Electric Cooperative, the Pacific Northwest Utilities Conference Committee, the Northwest Public Power Association, and the participation at various stages of other organizations in-

cluding the Solar Energy Association of Oregon and the Electric Power Research Institute. Thus, all the major "actors" in Pacific Northwest energy issues participated in the planning, design, operation, and evaluation of the Project.

The Project was a five year effort and cost \$20 million. The official Project had two basic areas. The "first experiment" was to retrofit homes. This was an experiment in administration, logistics, field organization, retrofit, and community participation designed to test a utility's ability to deliver services on a tight timeline. The "second experiment" was the research and supporting data collection for 16 evaluation and applied research topics. Guiding and uniting effort in these two areas was a "third experiment" in which, through creative use of theory and proceeding by trial and error, we evolved an organizationally intelligent working model for applied research. We believe this model is generally applicable. The "third experiment" was required to make it possible to conduct the weatherization and to evaluate the results.

From this perspective, the project was a test of the-

oretical assertions about the extension of quasi-experimental design developed by Dunn (1982), Campbell (1982), and others. This is an extension of technical design to optimize organizational factors that affect research work, and broadens the role and responsibility of the evaluation researcher. It aims to protect the evaluation effort against "threats" to the degree of validity and usefulness of results—protection against "Type III" errors in applied research.

Type I error is the assertion of a positive statistical conclusion that is actually false. Type II error is the failure to assert a statistical result that is actually present in the data. These errors are controlled through

sample design, selection of appropriate design formulas, and consideration of statistical power, statistical significance, and required precision. Type III errors include practical errors. The worst of these involve ways of misstating the research problem, inadequate and incorrect conceptualization, unconscious organizational or personal bias incorporated in the research process, doing a competent technical job solving the wrong research problem, or solving the right problem too late (Dunn, 1982). We report successes and problems with this approach, as experienced in the Project, drawing from recent work in the control of "Type III" errors and implementation of the Project's three experiments.

ORGANIZATIONAL FACTORS IN EVALUATIVE RESEARCH

We know that evaluative research is an organizational enterprise. Yet we need to become more creative and systematic regarding the organizational aspects of what we do. Consider, for example, the effort, technical emphasis, and allocation of resources used in guarding against statistical errors. Contrast this with the lack of effort, technical emphasis, and availability of tools to achieve excellence in structuring the organizational aspects of research. That we do not have systematic and well developed techniques for dealing with organizational factors is remarkable, because inadequate organizational structuring can invalidate work that is otherwise technically proficient.

Because organizational variables and the normative interests of organizations affect research results, they become the responsibility of the evaluator. Fortunately, there is a theoretical perspective that, when coupled with organizational techniques, can help evaluators negotiate the organizational normative pressures inherent in organizational settings so as to enhance the usefulness of evaluative research. We introduce this perspective, with guidelines for practice, in relation to concerns for research ethics, usefulness of results, and the goal of truth.

ETHICS

One basic kind of "Type III" error lies in the ethical dimension of evaluation practice. The error consists of being unaware of the social implications of technical choices, and/or of failure to orient technical choices toward "the good and true life."¹ In this perspective, ethical implications are inseparably embodied in the acts of technical choice. Ethics enters at the major stages of evaluation research, at the middle level of methodological preferences and choices, and also at the finer levels of technical decision.^{2,3} This perspective follows from understanding evaluative research as a sequence of choices with implications (probable consequences) to be acted out in the human world. Technical decisions have social consequences (while the fabric of organizational and social life both constrains and supports technical possibilities). Because technical choices have probable consequences, they are inherently ethical choices. This adds a burden of social responsibility to the work of the evaluator that may transcend, conflict with, or support organizational interests.⁴ Organizational cultures and contexts often direct our perception away from the inherent ethical dimensions of technical choices. But, supposing we accept the bigger picture in these everyday situations, how can we regain the ethical

dimension with enough clarity to accept responsibility in situations which initially present themselves as routine technical choice?

Some Guidelines

We propose the following questions for personal, organizational, and technical choices in evaluation activity, where "choice" refers both to the activity being evaluated and the evaluation itself:

1. Does this choice promote human freedom?
2. Does this choice promote "the good and true life" as the evaluator shapes herself/himself and as the project helps shape organizational and social relations?
3. Does this choice contribute to vitality, social learning, and well-being of people it affects?
4. Does this choice offer a reasonable chance of producing knowledge that will promote intelligence, social cooperation, efficiency, material welfare, and the potential to thrive?

In some situations the social implication of technical choice may be of no practical importance. But in many situations, even those portrayed as purely technical, social consequences are important. Here the evaluator

should show preference for "the good and true life." In other words, we look at evaluation work in the broad sense of human activity which contributes to the "creation" of the world in which we live and the conditions for life. This does not resolve questions of what to do in any simple sense, but it is important for defining the nature, quality, and contribution of our work that such questions as those above be raised in technical contexts.

Application

In the Hood River Conservation Project, all points of this criterion (ethics) were employed in the affirmative, contributing to the sense of exuberance and enjoyment we experienced in our technical work. In fact, in the field of conservation nearly all work appears adequate according to the above criterion because improved energy efficiency reduces the adverse environmental effects of energy production (greenhouse effect, acid rain, pollution), conserves scarce natural resources, and saves money for society and program participants. Hard work in developing practical knowledge about

how to "do more with less" is inherently self-actualizing and rewarding. For members of an evaluation team, the company of stimulating colleagues with "good faith" commitment is a source of growth, well-being, social health and personal refreshment. The orientation advocated is realistic in terms of the fundamental relations of humans and the ecosystem.

The long-term success of conservation efforts is keyed to empirical measurement and sound evaluation. This is the "apple pie" aspect of conservation with which our evaluation work is entwined. More deeply we touch here on the ground of being and the quality of wisdom in our search for knowledge. Our work in conservation evaluation is aimed at making it possible for people to live more efficiently, with fewer environmental insults, and with a more intelligently designed and least cost power infrastructure so that all may better thrive. Conservation and conservation evaluation enhance our survival prospects and our quality of life. This brings us to the dimension of usefulness.

USEFULNESS

We are interested in knowledge development guided by humane normative orientation.⁵ This raises the practical problem of producing useful results from the perspective of those parties with a stake in specific evaluations, and in particular from the perspectives of those host organizations that are the proximate environment and pay the bills for the evaluation and for the service delivery systems. If, under "ethics," we ask how to make our research useful to humans as a species, under "usefulness" we ask how to make our work useful within our specific organizational and interorganizational context. There are many activities that can enhance usefulness, but here we call attention to one of the most important: the underwriting of a climate of enthusiasm and of "free speech."

Some Guidelines

One formulation of a criterion for adequacy of evaluative research in this area is:

1. Is the internal climate of the research effort and the wider inter-organizational process guiding the research characterized by openness and inclusiveness? Are researchers free to speculate, develop and test ideas, use their own ideas and perspectives, express diversity, and work in a climate of personal growth?
2. Are both intra-organizational and inter-organizational processes oriented toward approximating the goal of the free speech community?

Evaluative research is conducted in the context of large-scale bureaucracies.⁶ On the level of appear-

ances, it is to the ends of these particular "structures-in-process" that the products of evaluation are to be of use. Yet there is a potential for contradiction in this arrangement, because of tension between the hierarchical character of large organizations and the goal of truth in evaluative research.

Organizational interests tend to be defined hierarchically, while the professional and technical staff is expected to implement lines of action and to act as if current organizational representations of reality are accurate. Yet such features of organizational life interact with the decisions of evaluation. Organizational commitments become embodied, often unreflectively, in practices and preferences of an organization's evaluation shop and in preferences in the selection of outside experts. Organizational tendencies toward homogeneity of perspective and technical preference are a source of difference in outcomes when similar lines of research are implemented by researchers studying similar problems in different organizational settings.

To counteract this tendency, "freedom as a criterion of truth" is a necessary technical condition of the validity of research.⁷ Further, the problem of research validity within an organizational context is one aspect of wider problem establishing a high degree of research validity in an inter-organizational social context. This problem can be addressed through "jurisprudence as an organizing approach" and the "transactional model of argument" (Dunn, 1982).

The Transactional Model

In "Reforms as Arguments," Dunn (1982) introduces jurisprudence as an organizing metaphor for applied

social research. His argument is that the metaphor of experiment, borrowed from the physical sciences, if used alone carries inappropriate meanings. Dunn proposes three levels in the conceptualization of applied social research: (a) the dominant metaphor of jurisprudence; (b) the metaphor of argumentation; and (c) the metaphor of the experiment. The work of Campbell and others in the development of the techniques and philosophy of quasi-experimental design is carried at the lowest level of conceptualization of the model.⁸

Dunn asserts five claims: (a) Campbell's metaphor of the "experimenting society" is insufficient because "Nature" does not cleanly edit the results of evaluative research.⁹ Instead, research results are "symbolically mediated" by the diverse standards, worldviews, frames of reference, problem definitions, ideologies, and commitments of policy makers, researchers, and other stakeholders in such research. (b) For this reason, attempts to solve practical problems (as in evaluative research) are better cast in the metaphor of argument rather than as experiments. (c) An appropriate model which "accommodates the experimental metaphor" and also accommodates the higher metaphor of argument is "a transactional metaphor of argument." The transactional model is an operationalization of the abstract metaphor of jurisprudence. (d) Under the metaphor of jurisprudence, tests of knowledge claims are referred to as "threats to usable knowledge" as they are in Campbell's earlier work on the design of quasi-experiments. (e) The transactional approach promotes freedom because it explicitly incorporates steps for surfacing "assumptions and implicit standards of assessment that shape and distort the production and use of knowledge" (Dunn, 1982, p. 296).

To extend the theory of quasi-experimental design, Dunn introduces "second order" threats to validity, which is the class of "Type III" error. Whereas the threats to validity established as technical criteria in the literature of quasi-experimentation involve threats in the dimensions of "internal, external, and statistical conclusion validity," and may be referred to as first order threats, second order threats (*i.e.*, Type III errors) call into question the definition of the research problem. This is exactly the problem in evaluation where worldviews, ideologies, and frames of reference are different.

In the metaphor of the experiment, we envision the development of knowledge through the competitive replication of key experiments, but in the perspective of the metaphor of jurisprudence as operationalized in the transactional model we recognize that replication is rare. More likely is the situation in which stakeholders engage in the "competitive replication of knowledge claims . . . in contrast to the competitive replication of experiments" (Dunn, 1982, p. 304) in research planning

and interpretation of results. In the light of these realities, how can our work be made most useful?

Application

The bottom line for usefulness of an evaluation is that a free speech approximation must be organizationally sustained. In the Project, the work of Campbell and others formed the technical basis for the quasi-experimental design. The work of Dunn (1982) and Campbell (1982) guided establishment of a free process and oriented the organizational structure and culture of the evaluation. The Project succeeded in approximating a free speech community. We employed the theoretical and technical perspective developed by Campbell and Dunn in the mode advocated by Restivo, that is as an "emancipatory epistemic strategy" (1983; Restivo & Loughlin, 1987).

For Hood River we report the following results (Peach, 1985b):

1. Continuing interest on the part of Regional Research Advisory Group (RRAG) participants. The RRAG was composed of representatives of several important energy-related groups in the region. High continuing interest in using Project results in regional policy forums.
2. Lively interaction in monthly meetings of RRAG, providing creative guidance throughout research planning, implementation, and results assessment.
3. The productivity of multiple perspectives in generating interpretations, amending research designs, and developing better research by anticipating counter-interpretations in the planning and implementation stages of the project.
4. Continuing mutual inspection and questioning of results within an underlying consensus format.
5. Widespread attention to the experiment in circles concerned with electric power and the environment.
6. Successful screening of many research "add-on" subprojects that might have diverted resources from primary project goals, through the give and take of the consensus process. Actual "add-on" research components had to survive intense questioning and negotiation.
7. Research open, with the characteristic power and focus of a free speech community, in observance of science norms and high technical standards; on time and within budget, due in large part to active oversight of interested parties.

We can also report conditions which appear required to make such a process possible (Peach, 1985b):

1. An initial basis for at least tentative respect among parties. Assembly of professionally competent and interested participants representing diverse views.

2. That none of the parties have the power to dominate any of the others (Peach, Oliver, & Goldstein 1984; cf: Lundy, 1984). That is, none of the parties' participation or freedom of perspective could be changed except through discussion and better argument.
3. Free flow of information, coupled with the right to independent inspection of work-in-progress and access to data for analysis.
4. Sufficient interest in the Project and/or need for continued cooperation.

A commitment to the goal of truth is one of the most basic of science community norms. It must often be defended in practice, with tact and diplomacy and sometimes by confrontation. This, again, is a part of the evaluator's and the evaluation manager's responsibility; skill in organizational diplomacy is quite important.

A Guideline

The criterion in this area is straightforward: Is the research effort consciously oriented toward the goal of truth?

Campbell emphasizes the importance of the goal of truth, and it is especially relevant when assumptions continually surface through an explicit consensus process. Campbell (1982, p. 329) warns that ". . . substituting the goal of persuasion for the goal of truth, or defining truth as consensus (amounts to) ontological nihilism." Yet there is no escaping the reality underlined by Dunn that a research program is essentially an argument, at all stages symbolically mediated. Campbell himself does not assert a simple correspondence theory of truth. Indeed, Campbell emphasizes the "indirectness, presumptiveness, and fallibility of all modes of knowing."¹⁰ Because human understanding of truth is always provisional and symbolically mediated, a critical and reflective orientation toward truth as a goal is critically important in applied research. This insight, the contingent nature of possible knowing, again underlines the arguments of the previous sections of this paper with regard to ethical criteria and the necessity to approximate in practice the free speech community.

Application

The criterion of the goal of truth as a test of the adequacy of evaluation was a reality in the Project. The

The explicit use of a consensus process employing the insights of the transactional model turns potential problems in the organizational relations of research into a positive force resulting in better research. Also, it is likely that only by such involvement will all parties continue to "own" research products when the results are in (cf: Dickey & Hampton, 1981). But if we acknowledge the transactional model and the role of multiple perspectives and consensus in developing useful knowledge, we must specifically integrate into this perspective the goal of truth.

TRUTH

participation of several organizations with overlapping but often disparate interests required consensus, with contingent episodes of data-referenced argumentation. The inter-organizational cooperation involved the participation of adequately funded and independent adversaries bound by the goal of truth. So long as the goal of truth was perceived to be operative, the parties retained their interest in cooperation. For this goal to be operative, the evaluation team and the RRAG were required to proceed under the approximation of free speech and to follow other science norms. The goal of truth stabilized inter-organizational cooperation and guaranteed the organizational conditions sustaining free inquiry and open research.

For example, two different interpretations were offered on the analysis of electricity use and savings for participants in the Project. The traditional evaluation perspective defined the Project's energy savings on the basis of the measured reduction in participant electricity bills, pre- versus post-Project. The planning perspective defined savings on the basis of the difference between a priori estimates of electricity use and subsequent energy-use levels. Initially, there was considerable disagreement about which perspective to present in the final report. After substantial discussion a consensus emerged that encompassed both perspectives.

A theoretical perspective organized by the conceptual touchstones of ethics, usefulness, and truth appears abstract, but it is actually quite material in its implementation and consequences. The Project provides an example of the productivity of such concerns when put into practice.

PRODUCTIVITY

The practical outcome of the Project's consensus approach was an experiment that worked. Not only did the "third experiment" work as anticipated, but integral to the success of the "third experiment" was the success

of the "first experiment" in the field and the "second experiment" in research and evaluation. Some features that emerged from this process were:

1. High quality research design (a nonequivalent control group design with two comparison communities and a regional random sample of homes).
2. A comprehensive Project evaluation designed before the Project was introduced in the field.
3. High quality instrumentation, including the development and implementation of a system to collect data on the energy output of wood stoves.
4. Use of a sociologist to perform a community assessment and completion of a baseline survey prior to the Project.
5. Continuing attention to the collection and management of relevant data.

This approach provided flexibility in redesigning aspects of the research based on practical experience as information became more complete. The process of cooperation allowed a close approximation to the ideal free speech community. It allowed the surfacing of assumptions at all points of design, implementation, analysis, and interpretation and supported the technical rigor of research throughout the Project. For example, as the fieldwork progressed, it became clear that wood use was a very important factor, confounding interpretation of electricity-use data. The RRAG decided to conduct a survey among participants to identify wood-use practices in Hood River. The spirit of professional cooperation among RRAG made it possible to modify Project activities as the need arose.

Drawbacks

We are essentially proposing a participatory approach to evaluation. In the Project, this was manifest through establishment and successful operation of the RRAG, supplemented by parallel efforts in the research teams. But establishment of a consensus process entails costs, both financial and in the amount of time it takes to conduct a project. The monthly research meetings were attended by representatives of several organizations; the time for these people represents a substantial cost. When consensus could not be reached on some aspect

of research design, project implementation, data analysis, or interpretation of results, discussion might proceed for two to three months both informally between meetings and in the monthly meetings. Sometimes results were held up until data could be analyzed with different assumptions (selection of cases, treatment of outliers, etc.).

Several research reports were redrafted until consensus could be reached, while a single research team and a single sponsor might have proceeded much more efficiently (although with less consciousness and breadth of perspective). In addition, the process of consensus showed minor vulnerability to unilateral "stands" by particular organizations. At various times, some of the cooperators "held out" for a particular position on a particular decision. The maximum time to resolve these was three months. The most serious threats to consensus occurred when a "higher up" in one of the organizations ordered a participant to take a position. The "higher up" was acting within the framework of their organization. Luckily, the discussion, constructive spirit, and creative intelligence within the RRAG always developed a creative way out of such situations.

Finally, we note that the processes put in place by affirmative response to the criteria specified here do not guarantee truth. A consensus style of operation develops its own culture, different from the hierarchical style of the participating organizations but not necessarily pure and honest. Each participant sees truth through particular lenses structured in part by organizational and personal perspectives, and it may take considerable discussion before we understand the elements of truth at the core of other people's perspectives. Another potential drawback is that every group tends to develop a hierarchy, and while "rank" may be pulled in overtly hierarchical organizations, in the consensus form of social organization articulateness and forceful enthusiasm may carry the day. Still, we can only pursue the goal of truth. We believe that the approach advocated here is a long term best bet for evaluative research.

SUMMARY

We propose a theoretical perspective embodied in seven criteria as measures of adequacy in the organizational dimension of evaluative research. The first four are guidelines for individual technical decisions in the conduct of research. They are concerned with the orientation and social legitimation of evaluative research as a human enterprise. The last three are guidelines for structuring research. They are designed to unleash energies and promote support by social structuring that reverses organizational tendencies that might otherwise interfere with research in ways that would lessen the degree of validity and usefulness of results:

1. Does this choice promote human freedom?
2. Does this activity or this choice promote "the good and true life?"
3. Does this project or this choice contribute to vitality, social learning, and well-being of the people it affects?
4. Does this choice offer a good chance of producing knowledge that will promote intelligence, social cooperation, efficiency, material welfare, and the potential to thrive?
5. Is the internal climate of the research effort, and the wider inter-organizational process guiding the re-

- search open and inclusive? Do researchers feel free to speculate, develop and test ideas, own ideas and perspectives, express diversity, and work in a climate of growth and self-actualization?
6. Are both intra-organizational and inter-organizational

processes oriented toward approximating the goal of the free speech community?

7. Is the research effort consciously oriented toward the goal of truth?

END NOTES

1. This focus on the "good and true life" is developed by Jurgen Habermas, as cited by his translator, Thomas R. McCarthy (1975, p. xv).
2. Compare ethics and policy. "Webster . . . defines ethics as 'the principles of conduct governing an individual or a profession, i.e., standards of behavior.' Policy is defined as a 'projected program consisting of desired objectives and the means to achieve them.' Obviously, the two concepts are closely linked, since ethical norms help shape which objectives are desired, and which means are chosen to reach them" (Adamek & Marvin, 1987).
3. Major stages are ". . . deciding to evaluate, specifying the evaluation agent; performing evaluation; reporting results; utilizing results to manage or modify programs" (Heilman & Martin, 1986). Middle range evaluation performance activity includes ". . . theoretical formulation, research design, in the sense of the structure of time/group comparisons, data gathering procedures and data analysis" (Bloombaum, 1987). Micro level activity includes specification of qualitative and quantitative aspects of approach to the evaluation problem, setting statistical power in sample design, planning contrasts, and the like. As this list of levels and activities suggests, evaluation as a human activity is a larger and more interconnected effort than the aspects on which we usually focus. We actually, however, deal with this totality as we proceed with our work. The aspects of our work take on relevance in relation to this totality of interconnected human activity in organizational, inter-organizational, and wider social context.
4. The recent historical roots of this perspective lie in the work of the "Frankfurt School" of social research. The work of this school represents an attempt by German social scientists to understand and undermine the basis of Nazism and Stalinism. (See McCarthy, 1973; Horkeimer & Adorno, 1982; Schäfer, 1983; Peach, 1985a).
5. Since our research work is part of the self-formative process of future social relations and of the human species, the first question to ask about the usefulness of research is its orientation and conduct with regard to the kind of world being created.
6. We include small independent shops and freelancers in this perspective because their income usually derives from contractual relationships with large-scale government, business, or industrial organizations.
7. "Freedom is a criterion of truth" is a central insight of critical theory and the Frankfurt School. This perspective, joined with the principle of scope and depth are the basis for the objectivity of research: "The quality of the validity generated by any given community is a function of the degree to which the social interests of the community are general and diffuse rather than specific and focused" (Restivo & Loughlin, 1987). More specifically this is "Type B" objectivity. In contrast, "Type A" objectivity is associated with narrowly focused research oriented toward subjects about which new knowledge is inherently generalizable (Restivo, 1983, p. 150). The focus in this paper is on research with strong "Type B" dimensions. Evaluation research is typically of this sort, because it combines physical and social science and tends to be highly relational to different interests and stakeholders.
8. Donald Campbell, Professor Emeritus of Northwestern University and Professor at Lehigh University, is generally recognized as the originator of the primary intellectual and technical advances which form the basis of evaluative research.
9. Dunn's argument is stated more broadly in terms of social research, including evaluative research, policy research, market research, etc.
10. Campbell (1982, p. 335). Campbell asserts the traditional "fact versus value" dichotomy, and emphasizes truth as a goal. In this way, Campbell seeks to maintain an open process of discussion as a counterbalance to what he terms the "belief manipulation interest" characteristic of both business system and socialist power structures. In so doing, he is aware that such assertions contribute to an aura of objectivity which serves to legitimate applied research but at the same time may be used in behalf of such belief manipulation interests.

REFERENCES

- ADAMEK, R.J., & MARVIN, G.M. (1987, April). *Ethical and policy implications of applied sociologists' work*. Paper presented to the panel on social policy, meeting of the Pacific Sociological Association, Eugene, OR.
- BLOOMBAUM, M. (1987, April). *Influence of research design upon data analysis*. Paper presented to the panel on quantitative analysis, meeting of the Pacific Sociological Association, Eugene, OR.
- CAMPBELL, D.T. (1982). Experiments as arguments. *Knowledge: Creation, Diffusion, Utilization*, 3(3), 327-337.
- DICKEY, B., & HAMPTON, E. (1981). Effective problem-solving for evaluation utilization. *Knowledge: Creation, Diffusion, Utilization*, 2(3), 361-373.
- DUNN, W.N. (1982). Reforms as arguments. *Knowledge: Creation, Diffusion, Utilization*, 3(3), 293-326.
- HEILMAN, J.G., & MARTIN, D.L. (1986). On the organizational theory of evaluation: Economic and political constraints and the choice of an evaluation agent. *Administration & Society*, 18(3), 315-333.
- HIRST, E. (1987). *Cooperation and community conservation* (ORNL/CON-235). Oak Ridge, TN: Oak Ridge National Laboratory.
- HORKHEIMER, M., & ADORNO, T.W. (1982). *Dialectic of enlightenment* (translated by John Cumming; originally published as *Dialektik der Aufklaerung*, 1944). New York: The Continuum Publishing Company.

- LUNDY, T.S. (1984). National program planning: Philosophy and example. Reprinted in *Doing better: setting an agenda for the second decade. Proceedings of the ACEEE 1984 Summer Study on Energy Efficiency in Buildings* (Vol. K). Washington, DC: American Council for an Energy Efficient Economy K-169 to K-174.
- McCARTHY, T. (1975). Translator's introduction. In J. Habermas *Legitimation crisis*. Boston: Beacon Press.
- PEACH, H.G. (1985a, April). *Insights from critical theory in applied research*. Paper presented at the 56th Annual Meeting of the Pacific Sociological Association, Albuquerque, NM.
- PEACH, H.G. (1985b, October). *Argumentation in applied research*. Paper presented to the 10th Annual Meeting of the Society for Social Studies of Science, Rensselaer Polytechnic Institute, Troy, NY.
- PEACH, H.G., OLIVER, T.V., & GOLDSTEIN, D.B. (1984). Co-operation and diversity in a large-scale conservation research project. Reprinted in *Doing better: Setting an Agenda for the second decade. Proceedings from ACEEE 1984 Summer Study on Energy Efficiency in Buildings* (Vol. I). Washington, DC: American Council for an Energy Efficient Economy.
- RESTIVO, S. (1983). *The social relations of physics, mysticism, and mathematics* (Studies in Social Structure, Interests, and Ideas). Dordrecht, Boston & Lancaster: D. Reidel Publishing Company.
- RESTIVO, S. & LOUGHLIN, J. (1987). Critical sociology of science and scientific validity. *Knowledge: Creation, Diffusion, Utilization*, 8(3), 486-508.
- SCHÄFER, W. (1983). *Finalization in science* (Boston Studies in the Philosophy of Science, Vol. 77). Boston: D. Reidel.