Arch. & Comport. / Arch. & Behav., Vol. 9, no. 1, p. 85-98 (1993)

Environmental Design Evaluation as a Change Oriented Research Process

Gary H. Winkel
Environmental Psychology Program
City University of New York
33 W. 42nd Street
New York, N. Y. 10036
U.S.A.

Summary

Environmental design evaluation (EDE) is a broad term that involves the structured assessment of the physical and social characteristics of a setting in relation to the activities that may take place there as well as the social and psychological consequences of the setting's design and use. It is argued that, given the scope of the evaluation process and the intention to use the research to bring about environmental change, the conceptual basis of EDE is different from that of most person/environment research raising issues about the way in EDE is conducted and how its validity might be assessed. A central element of the discussion involves the necessity to recognize and respond to the different value perspectives of those who use the setting being evaluated. Approaches to accomplishing this goal are presented and discussed in relation to validity concerns.

Résumé

Le terme évaluation de l'aménagement de l'environnement construit (EAE) est très large; il implique une appréciation structurée des charactéristiques physiques et sociales de cet aménagement, ceci par rapport aux activités qui s'y déroulent comme aux conséquences sociales et psychologiques de son usage. L'auteur pense que, compte tenu de l'étendue du processus d'évaluation et du fait que la recherche doit être utilisée pour introduire des changements, l'EAE se base sur des concepts différents de ceux utilisés par d'autres recherches sur l'environnement et ses usagers. Ceci provoque des questions sur la manière dont l'EAE est menée et sa validité établie. Au centre de la discussion se situe le besoin d'identifier et de répondre aux différentes valeurs adoptées par ceux qui utilisent l'environnement évalué. L'article présente des approches permettant de parvenir à ce but; celles-ci sont discutées en termes de validité.

1. Introduction

The statement of aims found in the frontispiece of this journal clearly indicates the need for an interdisciplinary approach to research designed to understand the "...relationship between human beings and built environment". The subject areas considered relevant to the journal's intent range from environmental aesthetics, participation, decision making, symbolic interaction and spatial behaviour to personality and

the environment. While each of these issues could be considered alone in the context of the built environment, there is one aspect of built environment research that could, in principle, involve them all - built environment evaluation.

The evaluation of the built environment has also been called environmental design evaluation (EDE). EDE studies seem to fall naturally into two categories. The first of these involves work undertaken as a prelude to possible design changes. This is predesign research and typically occurs in a existing setting. The second type involves the assessment of designs that have been implemented. In the literature, the latter studies are usually referred to as post-occupancy evaluations (Preiser et al., 1988). Preiser et al. (1988, 3) define post-occupancy evaluations (POEs) as

> "... the process of evaluating buildings in a systematic and rigourous manner after they have been built and occupied for some time. POEs focus on building occupants and their needs, and thus they provide insights into the consequences of past design decisions and the resulting building performance."

Because of its broad scope, one might ask if environmental design evaluation should be considered as an extension and expansion of more circumscribed areas of person/environment research. I believe that most person/environment researchers think of EDE in this way regardless of whether they do evaluation research themselves. Using criteria that have been developed in areas of person/environment research that do not have an explicitly applied focus, EDE work is often considered methodologically flawed (i.e. having dubious validity in the experimental or quasi-experimental sense (Cook and Campbell, 1979)) and theoretically sterile. To many EDE practitioners, the body of knowledge regarding person/environment relationships offered by non applications oriented researchers is usually too limited in scope and generality to be of much utility in applied settings. Although the mutual criticisms seem balanced, EDE workers continue to operate defensively both because they accept the legitimacy of the methodological criticisms voiced by academic researchers and because they continue to believe that the theoretical perspectives developed by academic researchers will ultimately be of use in applied problems. As a consequence, I think that EDE has suffered historically from being seen as merely one aspect of person/environment studies rather than as a separate endeavour having its own conceptual problems.

My reason for asserting that the validity of EDE studies involves a set of criteria that extends beyond that employed in most person/environment research is based on the nature of evaluation itself. There are three interrelated issues in evaluation work that distinguish it from person/environment research generally. These involve: (a) the often explicit intention to utilize evaluation results to bring about change; and (b) the recognition that the possibility and direction of change must take into account differences in value perspectives among those in the setting; and (c) differential access to the resources necessary to bring about change in the setting being evaluated. These three issues have implications for the design and conduct of evaluation research as well as for the validity criteria that would be used to evaluate the evaluation process itself. I plan to consider these issues in the remainder of this paper.

2. Change as a Central Focus of Evaluation

One of the bases for a comparison between environmental design evaluation and person/environment research involves the intentions of the investigator. In the latter, the research intent usually involves theory generation and/or testing. In the former, on

the other hand, the central focus involves data generation presumed to be important to bring about change. Almost all EDE work of which I am aware is initiated because members of an organization detect a problem or problems of some sort and believe, to one degree or another, that change may be required. In other instances, EDE work is undertaken because an organization is planning to build something new and wishes to select what is considered best from among the alternatives available. The potentiality for change is implicit in each of the alternatives considered. Even in those cases in which an environment is to be assessed to see how it is working, there is an implicit mandate for possible change should the evaluation result in the identification of problems. It is hard to imagine an instance in which one would undertake an evaluation where the potential for change was not an issue.

In the evaluation of designed environments, change is not an abstract concept. The setting's socio-physical environment immediately provides the material context for possible change. An essential aspect of the context for change involves the identification of the participants and the conditions required to bring about change. In this sense, the EDE process is inherently political because of the unequal power relationships existing among the various groups that may be involved in the evaluation process. In the existing EDE literature, the person who commissions the evaluation, the people who use the setting being evaluated, the evaluator, and the designer are most frequently mentioned as creating the context for evaluation. But there are at least four other groups that may contribute to an evaluation's political context. These groups involve the policy bodies in which the client may have a central or peripheral role, the implementors of change, the managers of the resulting changes, and finally the investigator's colleagues in the discipline. It is unlikely that either the values or the assumptions about human nature held by representatives of these different groups will be congruent. As a consequence, the interaction between values and power plays a critical role in change oriented evaluation. For example, Rivlin and Wolfe (1985) provide a sobering account of the ways in which the beneficial effects of the physical design of a psychiatric hospital for children were totally undermined by institutional staff whose values regarding the treatment of the children were at variance with the values of those originally responsible for the treatment program that led to the hospital's physical design.

I would suggest then that success in the implementation of change depends most strongly on: (a) the identification of people who are committed to the use of evaluation information within the organization as well as those who are in a position to implement change; (b) the evaluator's ability to understand the values held by members of each of the groups identified in a particular setting; and, (c) to serve, if possible, as a facilitator of change when value conflicts arise.

These conditions have implications for the development of the evaluation team, the types of questions asked in evaluation, the research strategy to be employed, and validity problems in predesign and postoccupancy evaluation. The remainder of this paper focuses on each of these issues.

To make the discussion of conceptual issues in EDE work more illustrative, I will use a study conducted by Richard Olsen and myself on the redesign of an emergency medicine services (EMS) unit at a large urban hospital.

3. Project Description

The EMS actually consists of five divisions - the adult emergency services, emergency ward, pediatric emergency, psychiatric emergency, and walk-in clinic. The basic design problem involved the physical layout of each of the divisions and their relationships to one another. We were able to identify many areas in which design and policy changes could possibly have improved the overcrowded, confusing, and stressful conditions seen in the various services.

Although we were to conduct the evaluation research, we wanted people who worked in the EMS to be involved directly with us in the development of the issues to be considered and the questions to be asked. We did this because in an extensive study of factors in program evaluations that led to their use in bringing about change, Patton (1978) found that two variables were consistently identified as most important.

The first of these involved the political context of the evaluation (usually involving conflicts regarding budget allocations necessary for implementation) and second, the presence of an individual or group who "personally cared about the evaluation and the information it generated" (Patton, 1978, 64-70). Of the two issues, the latter was generally identified as most important.

At the beginning stages of evaluation, the investigator must locate those people who want to know something and who are in a position to use the information that might be generated from the evaluation. These people must have an interest in participating in the research process itself to ensure that the information generated is relevant to the needs of those who wish to act upon it (Greene, 1987, 1988; McLemore & Neumann, 1987; Patton, 1978). As Patton (1978, 72) points out:

> "What fundamentally distinguishes the utilization-focused approach to evaluation from other approaches is that the evaluator does not alone carry this burden for making choices about the nature, purpose, content, and methods of evaluation. These decisions are shared by an identifiable and organized group of decision makers and information users."

This point seems rather obvious. But I think we are all aware of the many instances in which the evaluator becomes the *de facto* decision maker either

> "...by default (no one else is willing to do it), by intimidation (clearly the evaluator is expert), or simply by failing to think about or seek alternatives" (Patton, 1978, 72).

In the EMS project, we made efforts to ensure that the two people most interested in the evaluation, the Director of Emergency Medicine and the Administrative Nurse for Emergency Services, were included in the evaluation planning and they provided invaluable assistance. I believe that the desire on the part of the EMS staff to bring about change using our evaluation findings underscores the points made by Patton and others. Participation in evaluation planning by those who wish to use the evaluation is essential. But it is equally necessary to ensure that *one or more persons in the evaluation group has access to, or can themselves, control access to the resources necessary to bring about change.* To the extent that this condition is not met, the chances of successful implementation are seriously compromised.

In recognition of the importance of this point, we had also secured the approval and encouragement of the Hospital Director who believed that there were problems in

the Emergency Services and who indicated her willingness to implement change (having done so on other projects we had undertaken at the hospital).

Once having established the evaluation planning group, we next turned to the problem of determining the range of people whose environmental assessments are to be included as part of the evaluation effort. In short, how representative will information gathering be? Most EDE workers incline toward the maximum (depending in part on budget considerations, of course) for two reasons. First, the desire for maximum possible participation stems from the belief that if we want to understand how environments "work", the experiences of all those who have direct involvement with the setting constitute the best data source. Second, maximum participation expands the range of those who feel they have a stake in the outcome of the evaluation effort. To the extent that these stakeholders believe that change will be of benefit to them, they have a greater incentive to press for the implementation of change. In this project, all EMS nursing staff as well as a ten percent sample of patients (who could be and were willing to be interviewed) and visitors were to be included in data gathering.

In addition to determining the range of people who are to be involved in information gathering, it is important to consider carefully how information from evaluation participants is to be used. Resolution of this question should lead to the development of a contract between the evaluation group and those who will provide needed information about the setting's operation.

The evaluation group, however, must determine the nature of the contract that exists between it and those whose assessments are to be used. Under ideal circumstances, the contract, usually stated verbally, should specify what kind of information is to be gathered, from whom, and how that information might be used to make recommendations for change. As part of this process, it is my belief that study participants should be made aware that different points of view about the setting and how it might be changed will be expressed. Furthermore, since not all proposals for change are feasible, any recommendations are subject to further negotiation.

The development of a contract in applied social research is clearly a value laden decision. There have been only occasional discussions of this issue in the EDE literature. For example, in a somewhat scathing critique of evaluation research, Knight and Campbell (1980) indicate that most evaluators limit themselves to the client's definition of the parameters relevant to design evaluation. As a consequence, those who use the setting but are powerless to control it are exploited. These authors argue that since evaluation is value laden, the evaluator should aim for the development of two or more sets of criteria that reveal the different values and perspectives of those who use an environment. They suggest that the evaluator should not simply serve as an agent of the *status quo*. Rather he/she should be an advocate for organizational change that may be counternormative. Most of their examples are drawn from problems arising from the design of work environments. They are critical of efforts that are made to maximize the technical efficiency of work environments because this is usually accomplished at the expense of those working in these settings.

Clearly, possibilities for exploitation exist in many organizations and the evaluator must decide personally how he or she will handle them. Shibley and Schneekloth (1985) and Schneekloth (1987) discuss some of these dilemmas. The issues raised by these authors are clearly relevant to the relationship between the evaluation group and the users of the setting under assessment. On the other hand, there is a certain insularity in these discussions. EDE workers are not the only people who confront these

issues. Those working within organizational frameworks are well aware of the problematic relationships that stem from power differentials. My reading of the organizational literature strongly suggests that there is a growing awareness that exploitation as a way of organizational life is, in the long run, generally counterproductive for everyone involved (Katz & Kahn, 1978; Schein, 1969, 1980). The communication of findings from this literature obviously does not guarantee that decision makers in the evaluation group will necessarily be sympathetic to the broader range of alternatives for change that might grow out the experiences of those who use the environment being assessed.

From my experience, it is rare to encounter much opposition to the broadest possible inclusion of setting users since it is these people who are experiencing the problems. Should opposition to broad participation occur, I usually argue that evaluation efforts that do not include setting users may very well be counterproductive. In large part, this is because of the close relationship between physical and social change that can occur at a number of different levels. It is therefore counterproductive to ignore the potential social consequences that result from changes in the physical environment.

In the EMS project, our verbal contract with those who participated in the study clearly stated that we, as evaluators, were in no position to know to what extent and how the information being gathered would be used. We indicated, however, that there was commitment to change from the Hospital Director.

The necessity for thinking carefully about the nature of the contract between those responsible for conducting the evaluation and those whose information will be used in making change recommendations is clearly important as an ethical issue in evaluation. The need for a contract, however, has implications for the framing of research questions.

4. Evaluation Questions

Among the reasons why evaluation efforts fail is that the evaluator does not clearly identify the types of questions for which answers are sought nor finds the users and/or decision makers for whom the questions are important. Smith (1987) indicates that there are four different kinds of questions that characterize evaluation research - those dealing with research, policy, evaluation, and management issues. Because EDE is essentially applications oriented, each type of question and the way in which it might be answered are problematically linked to one another as a function of the values and the differential power structures that characterize the positions of researchers, policy makers, managers, and implementors of possible change.

Unfortunately, in the published literature, discussions of the complex interrelationships that exist among the different types of questions that typically figure in any evaluation are meagre. To a large extent, this situation exists because of the emphasis placed on research questions to the relative exclusion of other types of questions that might be posed. For example, evaluation reports are typically written to include: the setting's description, a list of "problems", a discussion of the methods used to "clarify" the "problems", a set of design alternatives that might solve the "problems", and a report of the extent to which project recommendations were implemented. The rationality conveyed by a typical report's organization belies the convoluted and chaotic rationality that characterizes the evaluation process itself.

Consider, for example, the question of "problem identification". In its broadest sense, the term should involve all the types of questions identified by Smith (1987). In practice, as well as in most texts on research methods, however, problem identification is usually confined to the development of research questions. It is implicitly assumed that the provision of the "correct" information is sufficient to implement change. Bechtel (1988), however, notes that the empirical evidence supporting that assumption is almost non-existent.

Once having decided upon the group of people who will be involved in the evaluation, preliminary interviewing designed to elicit information that will lead to research, policy, evaluation, and management questions should be organized around understanding the operations or processes that are characteristic of the setting under study, the impacts of these operations on the various groups of people who are involved in the setting, and wherever possible, the different values that people employ when commenting upon both operations and impacts.

Given the multiple value perspectives that characterize most evaluation settings, particular attention should be given to the following questions: (a) *why* is an issue a problem (an evaluation question); (b) *for whom* is an issue a problem (a combination of an evaluation and research question based on the evaluation group's decision regarding the range of people whose assessments are to be considered part of the evaluation); (c) how do issues relate to one another in the ongoing functioning of a setting (a research and evaluation question); and (d) is the issue a problem or can it be made to be a problem for the people who must implement change (a policy question)?

It is important to recognize that the issues identified as part of the evaluation procedure only become problems that might be addressed under three conditions. First, there must be a perceived mismatch between the workings of a particular part of the environment and the values regarding the setting's appropriate operation held by those who experience it. It should be noted that different people will most likely have different views regarding the appropriateness of a setting's operations.

Second, a consensus must be developed among the evaluation team members regarding the importance of the mismatches considered candidates for change. Third, the problem must have a potentiality for change that is probable, not just possible.

The single factor having the greatest influence on the probability of change involves the relationship between policy questions and the existence of incongruent values held by people with different amounts of power within the organization itself. For example, in our preliminary reconnaissance of the Emergency Services, we discovered that psychiatric patients brought to the emergency room experienced exceptionally long waits in very uncomfortable surroundings before being seen by a physician. In an effort to understand why that was the case, our interviews with both adult emergency services and psychiatric nursing staff revealed that the EMS Director had made it a matter of policy that all emergency psychiatric admissions had to be given a physical examination by regular emergency room staff before being seen by a psychiatrist who would make a judgement regarding admission to a psychiatric unit. In terms of medical emergencies, however, it was clear that the examination of psychiatric patients was assigned a low priority for the overworked emergency room staff. The line nurses were aware that this policy created multiple problems for psychiatric patients. They also indicated to us that they had argued strongly that psychiatric staff should be responsible for physical examinations. In this instance, we felt that good design would not cure bad policy. In discussions with the EMS Director, we pointed out that, given the

mental condition of most emergency psychiatric admissions (apathetic withdrawal and/or agitated confusion), design changes would probably accomplish little in making the wait more pleasant. Our research, however, documented the nature of the problems created by the policy decision.

5. Organizing the Research Process

The development of a evaluation research strategy depends on the nature of the EDE project. In predesign work, the intent is to generate an understanding of the ongoing flow of activities within the setting to be evaluated. Among other things, how do the functional activities of the setting operate and how do they fit together in the physical space provided for these activities?

Prior to this phase of the study, we had walked through each of the EMS divisions with the Medical Director and Nursing Services Director noting, from their perspectives, where problems existed with regard to the delivery of treatment and began to identify how, why, and for whom these problems had an impact. Using the same approach, we conducted interviews with physicians and nursing staff in each of the divisions. These steps provided an overview of EMS operations, an understanding of differing views regarding problems and their impacts on staff, patients, and visitors.

Using this information, we were next able to organize our observations of the daily activities within each of the EMS divisions. For example, we followed the progression of a sample of patients from time of entry through to transfer to another hospital unit or their discharge. During this period, we wished to document where they spent time, with whom they interacted, and how the treatment process progressed. At the same time, we noted how these activities impacted on other functions. We then used our observational data to design further interview questions to clarify the treatment and management policies that would help us understand the meaning of what we observed.

This approach to evaluation research is designed to understand the setting's processes and context as an ensemble of activities as well as differential evaluations (for example, the "goodness-badness", helpfulness-harmfulness, appropriateness-inappropriateness) of the consequences associated with the fits and misfits among component processes by different setting participants (for example, different perspectives on waiting for treatment or interactions with nursing staff as assessed by the patients and nurses who may have different perspectives and values regarding the ensemble's operation as a whole).

This ensemble oriented approach to evaluation does not preclude a focus on component processes and values. The meaning of any component, however, can only be understood in the context of the whole. An interesting illustration of the whole-part relationship in our EMS work once again involved waiting for treatment.

The site of our EMS research is a municipal hospital and it serves the medical needs of prisoners from the city's prison system. Prisoners requiring treatment may be seen in the EMS. They are accompanied by police officers and are manacled. This situation created a very real conflict for the nursing staff. As providers of medical care, their general value regarding the right to treatment regardless of a patient's personal circumstances justified the use of the EMS for this purpose. At the same time, however, they felt that in the context of the special anxieties associated with emergency medical treatment, exposing other patients and visitors to people who are

potential threats would lead to greater tension in the waiting room. This conflict arose because the nurses assumed that patients and visitors both did not share their more general values regarding treatment access and felt that the patients blamed the staff for exposing them to prisoners.

In an effort to allow for the expression of patient's values regarding this issue we designed our research to include both open-ended questions about the presence of police and prisoners and we also had some scaled questions regarding the degree to which people felt upset during the waiting period. Because police and prisoners were only intermittently found in the waiting areas, we had a kind of natural experimental situation. Somewhat surprisingly, we found that while patients reported being more upset during those periods when police were present (controlling for other variables like length of wait and density levels in the waiting area), 55% of the patients preferred police presence because the officers could maintain order should problems arise. In addition, those who were upset by the presence of prisoners nonetheless felt that prisoners had a personal right to treatment and appreciated that the EMS provided it.

Our finding regarding the desirability of police presence was useful because it underscored the underlying degree of tension that characterized the waiting experience and provided more tangible evidence of the complexities involved in redesigning the waiting area. Perhaps more important, though, the nurses realized that while their concerns about the waiting process for patients and visitors were justified, the problems should not be attributed to their broader views about the EMS as a setting for egalitarian treatment.

This example underscores the importance of an approach to evaluation research that emphasizes the identification of different value perspectives while maintaining a concern for a setting's part-whole operation. At the same time, it raises questions about the validity of the research findings and the role that validity issues play in change oriented evaluation.

6. Validity in Change Oriented Evaluation Predesign Research

In predesign research, validity essentially involves an indication of the *fidelity* with which the evaluation team has been able to identify the entire range of problem sets that are relevant to each group of participants who have provided evaluation information and, to the extent possible, has been able to link these problem areas to observed uses of the setting. The main criterion for this aspect of validity involves agreement among setting participants that their problems have been identified and that, where relevant, their observed use or their report of how they would use the environment reflects the problem under consideration. The linkage between verbal report and observational data provides grounding for the evaluation group's understanding of setting issues.

Both verbal report and observational strategies have their own validity issues. Since verbal reports constitute a measuring device, it might seem appropriate to use the classical definition of validity - a measure's validity depends on whether it actually measures what the investigator wants it to measure. In evaluation research, however, there is a major difference. There is no single investigator who makes the decision regarding the validity of any measure. The process of deciding what is to be measured (that is, the problems in the setting) is usually a negotiated process involving both the evaluation group and the setting's participants.

The resulting problem statements are based on three assumptions about values and their relationship to beliefs and actions within a setting. First, from the range of activities that either occur or do not occur in the setting being evaluated, each person selects out those that result in human consequences that are incongruent with their own values regarding how the setting *should* work and, given that, what are acceptable and unacceptable consequences. Second, the setting's program constrains people's views of possibilities for action in the setting. And third, it is assumed that these constrained values translate into experiences of, and action in, the setting. From these assumptions it follows that if there are multiple beliefs regarding what is desirable in an setting (the value issue) and these translate into multiple experiences and actions, (constrained, of course, by the rigidity of the setting's program) *no single validity assessment is either desirable or possible.* There are multiple (but not infinite) validities which are constantly being negotiated within the constraints set by the setting's program. If the evaluation is to lead to change there needs to be a consensus among setting participants about the legitimacy of the range of problems identified. Issues that are to be considered as problems and the priorities attached to their resolution is usually a matter of negotiation. For example, the triage nurses in the emergency services expressed greatest concern about waiting room tensions and patient privacy issues. Staff nurses in the Adult Emergency Service, on the other hand, were more concerned about the amount of time they had to spend walking to patients because of that department's layout. Both groups of nurses considered each other's problems legitimate even though their priorities differed.

This example would then suggest that a central aspect of validity assessment in predesign evaluation studies requires that information gathered about the setting's operations and the values held by setting participants be *representative.* In the EMS project, while physicians may have more power in the organizational hierarchy, this does not mean that the nurses' perspectives are irrelevant to the system's operation. Clearly, conclusions about the EMS based only on the physicians' perspectives would be less valid and credible than using a representative sample of those who use the setting.

In addition to the representativeness of the information gathered, the validity of research findings can be improved through greater attention to the *contextualization* of setting experience. In the EMS study, we used too many close-ended questions in the mistaken belief that numerical information would improve the validity of our results. For example, when nurses told us that people became upset and annoyed as a consequence of their waits, we prematurely accepted that assessment and included scaled items in our questionnaire regarding how upset and annoyed people were without determining the contexts in which people did or did not experience these feelings. While it seemed plausible that people might be both upset and annoyed, our results could have led to a more diversified understanding about these issues had we spent more time trying to understand why patients might be upset.

An anomalous finding based on the numerical data underscores this point. We learned that *actual* waiting times had little to do with the degree of upset people experienced. Rather their upset was more closely associated with how upset they felt generally and how long they *perceived* the wait to be. We learned almost nothing about factors that may have influenced perceived waiting time since we, as well as the line staff, were convinced that actual waiting time was the "real" problem.

Once the initial verbal reports about the setting have been generated, observational procedures need to be developed. Observational data involves an effort to trans-

late the multiple realities revealed in the verbal reports into the integrated activities that can be observed in the setting. It is assumed that the activities that are observed have meaning to the participants and the analysis of the observational data involves examining these meaning-activity linkages in the context of the setting's program and people's value systems. For example, the triage counter allowed few opportunities for the nurse to have a private conversation with a patient. Because of their concerns about patient privacy, the nurses established their own rules about people waiting to be triaged. The person behind the patient being interviewed had to stand at least four feet away from the triage desk, even though at times that meant that patient lines interfered with circulation through the entry corridor to the emergency room. We also saw various instances of the *absence* of potential activity-meaning linkages. For example, concerns about patient privacy were also expressed by the nurses in the emergency ward. Each bed in the ward had a movable curtain which could be used to provide at least visual privacy. However, rarely did we see these curtains being used.

The process of interpreting the meaning of behavioural observations usually involves gathering further information from the setting's participants either to enquire about the meaning of the observed or to ensure that the interpretation(s) of the observed behaviours are correct.

Once the evaluation group believes that the data they have gathered are faithful to the setting's operation, evaluation results can be used for the development of both design, and perhaps, policy changes. The evaluation of the proposed plan for change, in turn, needs to be assessed in terms of its abilities to accommodate the needs of those whose data generated the evaluation. In general, the greater the range of problems accommodated in the proposed change, the wider the group of people who feel they have a stake in the outcome. As a result, these people will be more likely to focus their efforts to bring about change. In the EMS study, our final report achieved this objective.

Underscoring the importance of having someone within the organization committed to change and capable of making resources available for change, our project recommendations were not implemented for the following reason. During the course of the evaluation effort, the Director hired an Assistant Director whom she placed in charge of the EMS work. That turned out to be disastrous for the ultimate project because after the final report was submitted he evaded our efforts to discuss implementation. Two months later, the Assistant Director was dismissed for allegedly misappropriating hospital patient funds. The Director was also dismissed because of her failure to conduct an adequate search of the Assistant Director's past employment records. The EMS staff, however, are using the report to make whatever small scale changes they can to improve the quality of EMS services.

7. Validity in Postdesign Evaluation

Assuming that the evaluation group has been able to generate a faithful description of the predesign setting and this description has been used to develop a changed environment, the type of research design to be used in the postoccupancy evaluation is still problematic. As Wener (1989) notes, the number of POES utilizing truly comparative designs is small and requires a degree of research support which is rare in social science circles. The majority of POES with which I am familiar frequently do not provide information regarding a setting's operation prior to the implementation of

the new design since the investigator does not have access to those data. The interpretive difficulties confronting those wishing to use some form of quasi-experimental design under these circumstances are well documented (Cook and Campbell, 1979) and need not be repeated here.

Even in those situations in which a comparative research design is feasible, both Wener (1989) and I (Winkel, 1987) have expressed misgivings about the descriptions provided for the different settings used for comparison purposes. Not the least of these involves the minimal number of instruments available for the description of the physical environment itself. As Wener (1989) indicates, generic terms like "open offices" and "low income housing" do not refer to a single type of setting but to a number of variants possibly grouped under one common theme. While Wener (1989) reviews some encouraging attempts to create instruments for environmental descriptions, this aspect of EDE work is clearly underdeveloped at present and creates both theoretical and practical obstacles to the determination of which of many aspects of the changed environment are responsible for observed changes in attitudes and behaviour.

In light of the problems of comparative research design, Wener's (1989) contribution to EDE methodology involves his recognition that there is a place for evaluation research designed to generate a descriptive understanding of the changes that have occurred in a new setting's operation even in the absence of comparative data.

8. Conclusions

My reflections on some of the conceptual issues that are important in EDE work suggest that greater attention in EDE work should be paid to the interrelationships between values and change in organizational settings. This interest also reflects a relatively neglected area of environmental research - an understanding of how values regarding the appropriate operation of settings are translated into environmental assessments and action by the many people who participate in them. Perhaps this paper will provide the necessary motivation.

BIBLIOGRAPHY

BECHTEL, R. (1985), Onward, upward and round about: to the future of the POE, *J. Architectural and Planning Research*, V (1988), 339-358.

COOK, T. & CAMPBELL, D. (1979), "Quasi-experimentation: Design and Analysis Issues for Field Settings" (Rand McNally,Chicago).

GREENE, J. (1988), Communication of results and utilization in participatory program evaluation. *Evaluation and Program Planning*, XI (1988), 341-351.

GREENE, J. (1987), Justifying conclusions in naturalistic evaluations. *Evaluation and Program Planning*, X (1987), 325-333.

KATZ, D., & KAHN, R. (1978), "The Social Psychology of Organizations" (John Wiley & Sons, New York, 2nd Ed.).

KNIGHT, R., & CAMPBELL, D. (1980), Environmental evaluation research: evaluator roles and inherent social commitments.*Environment and Behavior*, XII (1980), 520-532.

MCLEMORE, J., & NEUMANN, J. (1987), The inherently political nature of program evaluators and evaluation research.*Evaluation and Program Planning*, X (1987), 83-93.

PATTON, M. (1978), "Utilization-focused evaluation" (Sage Publications, Beverly Hills, CA.).

PREISER, W.; RABINOWITZ, H. and WHITE, E. (1988), "Post-occupancy Evaluation" (Van Nostrand, New York).

RIVLIN, L., & WOLFE, M. (1985), "Institutional Settings in Children's Lives" (John Wiley & Sons, New York).

SCHEIN, E. (1980), "Organizational Psychology" (Prentice-Hall, Englewood Cliffs, NJ., 3rd Edition).

SCHEIN, E. (1969), "Process Consultation" (Addison Wesley, Reading, MA.).

SCHNEEKLOTH, L. (1987), Advances in Professional Applications, *Advances in Environment, Behavior, and Design*, Vol. 1 (Zube, E. & Moore, G., Eds.), (Plenum Press, New York.

SHIBLEY, R., & SCHNEEKLOTH, L. (1985), Risking collaboration: Professional dilemmas in evaluation and design. *Journal of Architectural and Planning Research*, V (1985), 304-320.

SMITH, N. (1987), Toward the justification of claims in evaluation research, *Evaluation and Program Planning*, X (1987), 309.

WINKEL, G. (1987), Implications of Environmental Context for Validity Assessments, *Handbook of Environmental Psychology* (Altman, I., & Stokols, D., Eds) (John Wiley & Sons, New York, Vol.1).

WENER, R. (1989), Advances in the Evaluation of the Built Environment, *Advances in Environment, Behavior, and Design*, Vol. 2 (Zube, E., & Moore, G., Eds.) (Plenum, New York).