

THE EVALUATION OF PHYSICAL SETTINGS:  
TO MEASURE ATTITUDES, BEHAVIOR, OR BOTH?

Arthur H. Patterson  
Assistant Professor

Romedi Passini  
Research Assistant  
Division of Man-Environment Relations  
The Pennsylvania State University

A recurring problem in the evaluation of physical settings is that the dependent variables (what is measured) are usually insufficient to provide valid and generalizable information. Numerous evaluation studies are reviewed, and shown to utilize either attitudinal (e.g., satisfaction) or behavioral (e.g., use) measures. This paper explicates the threats to validity in using either attitudes or behavior individually. Further, both behavioral and attitudinal measures are defined in the context of evaluating designs, and their strengths and weaknesses are presented. Finally, a multi-method technique of measurement is proposed, where behavioral measures are used to validate attitudinal measures. Only in this way can sound and valid data be obtained.

Many evaluations of physical settings fail to provide data of value to the designer for a simple reason: the dependent variables, that is, those things that are measured about human behavior in the setting, are often insufficient to assess the designs effect. The purpose of this paper is to show how both attitudes and behavior must be measured in order to provide useful data.

Bases of Comparison. In order to do this, we must first give a very general view of what we consider the nature of evaluation to be. Evaluation in a literal sense might be understood as an act of establishing a value index for an element; at the base of evaluation lies a process of comparison of that element with others. Different bases for comparison can be ascertained from the research literature. These are:

- 1) Physical settings can be evaluated by comparing them to settings of similar content and function. This approach was used by the "Building Performance Research Unit" to assess school buildings in the United Kingdom. The base for comparison being, among other things, a performance profile of the building called a psarchigraph. (Markus, 1972).
- 2) Physical settings can be evaluated by comparing them to settings of different order occupied or known by the same users. In this case, the base line for comparison is the users' experience. (Canter, 1972). In a similar perspective physical settings can be evaluated by comparing them in a less controlled fashion to the experience people have with buildings in general. Overall satisfaction ratings might be an example of such an approach. (Canter, 1968).

3) Comparison might be undertaken in a more experimentally controlled fashion. Aspects of physical settings can be assessed by comparing settings that vary on the researchable aspect alone. These aspects might be, for example, of a physical, behavioral, or demographic nature. (Ittelson, 1970).

These processes of comparison can be used, as indicated above, to establish a value index for the building. To find this index though is rarely the aim of research. However, for the purposes of this paper, evaluation is seen as a means, rather than an end in itself. What is important is that in all three types of comparisons listed above, what is measured is some aspect of human experience: either attitudes or behavior.

Types of Dependent Variables. The three types of dependent variables most commonly used in the building evaluation literature are:

1) Behavior. Units of overt behavior might be regrouped into two categories: (a) units that refer to episodes of behavior on a small scale such as arm movements, head movements, etc., which are identified as molecular behavior, and, (b) units that refer to behavior of a more general nature composed of sets of molecular behavior and comprising activities like eating, working, etc., which are identified as molar behavior. We are mainly interested in molar behavior (at the building and extra building scale), molecular behavior being essentially the concern of ergonomics.

2) Attitudes and Cognitions. The cognitive units refer to the meaning or message that is received from the setting. The environment is seen as containing information that is extracted and interpreted by the occupants. The attitudinal unit usually refers to some overall reaction to the setting or parts of it. In building evaluation, one of the commonly used attitudinal units is in the form of an overall measure of satisfaction.

3) Demographic characteristics of the user. This variable is seen as a basis to develop a user taxonomy describing differential interests, attitudes, and behavioral habits of identifiable population sub-groups. As can be seen, the measures used here are actually those of (1) and (2) above.

It should be noted that the variables we have described have been commonly grouped together and called "behavioral issues". As we shall present below, attitudes are quite different than behavior, and should be recognized as such. Table 1 summarizes the main behavioral units of analysis as they have been discussed in the text. Each unit gives examples of illustrative research projects, a short description, and the necessary references. The table should not be viewed as an exhaustive review, but rather as illustrative examples.

Table 1. An Illustrative Review

EVALUATION OF SETTINGS ON THE BASIS OF:	ILLUSTRATION BY AUTHORS	STUDIED SETTING	RESEARCH DESIGN DATA GATHERING	TYPE OF RESEARCH RESULTS
BEHAVIOR DESCRIPTION -type of behavior -attribute to behavior efficiency dimension quality dimension	Esser        70	Ward	Case study guided observation	Territorial behavior in function of personality characteristics
	Bechtel     72	Residence	Comparative case studies behavior setting method	Comparison of behavioral characteristics (attributes) for two housing developments
	Durlak      72	School	Comparative case studies guided observation	Effect of open class layout on behavior (efficiency attribute)
	Hahn        73	Treatment Center	Case study behavior mapping	Occurrence and location of desired behavior as defined by treatment objectives
BEHAVIOR DENSITY number of behavior units in a setting	Ittelson    70	Hospital	Comparative case studies behavior mapping	Comparison of different hospitals in terms of behavior density
	Bechtel     72	Residence	Comparative case studies behavior setting method	Behavior density as a measure of the behavioral richness of a setting
BEHAVIOR DISTRIBUTION distribution of behavior in space including movement through space	King        69	na	na	Point pattern analysis, description and analysis of distribution pattern

Table 1 (continued)

EVALUATION OF SETTINGS ON THE BASIS OF:	ILLUSTRATION BY AUTHORS	STUDIED SETTING	RESEARCH DESIGN DATA GATHERING	TYPE OF RESEARCH RESULTS
	Bechtel 70	Museum	Experiment mech. device unobtrusive measure	Traffic pattern, movement and distribution of people in given space
	Rivlin 73	Ward	Comparative case studies behavior mapping	Age related differences in the use of the ward in terms of behavior distribution
BEHAVIOR STRUCTURE -functional link between behavioral units -behavioral compatibility -profiles of activities	Ittelson 70	Hospital	Comparative case studies behavior mapping	Activity profiles in hospitals
	Akin 73	University Campus	Case study free observation questionnaire diary	Misfit between physical structure and behavioral structure
ATTITUDE - SATISFACTION measure of overall satisfaction	Davis 70	Residences	Survey questionnaire	Identification of physical factors having the strongest effect on user satisfaction
	Canter 72	Hospital	Case study questionnaire	Overall and detailed appraisal of new hospital on the basis of satisfaction ratings
	Markus 72	Schools	Comparative case studies questionnaire	General evaluation of schools on a behavioral, economical and structural basis Description of an evaluation program

Table 1 (continued)

EVALUATION OF SETTINGS ON THE BASIS OF:	ILLUSTRATION BY AUTHORS	STUDIED SETTING	RESEARCH DESIGN DATA GATHERING	TYPE OF RESEARCH RESULTS
OTHER ATTITUDINAL MEASURES attitudinal dimension	Canter      68	na	na	Description of a cognitive approach to building evaluation
	Denton      73	Residences	Survey questionnaire	Development of a user taxonomy for university residences
MEANING communication potential of the setting	Wools      70	Room	Experiment simulation by drawings	Study indicating an effect of the meaning that a setting has on the observer's behavior

Purpose. As can be seen from this representative review, most of the studies do not combine both behavioral and attitudinal measures. The purpose of this paper is to argue that used individually, both behavior measures and attitudinal measures are insufficient to provide valid and generalizable data. However, when combined in any given evaluation, this multi-method approach results in rich, valid, and generalizable information.

Problems With Attitudinal Data. The reasons why attitudinal data (such as surveys or questionnaires) alone are insufficient has been dealt with at length elsewhere. Patterson, writing on the utility of unobtrusive measures (1974), and Wicker, writing on attitude-behavior inconsistency (1969) are good examples. To briefly summarize here, the major problems with attitudinal measures are:

- 1) Attitude measures are generally "reactive". That is, they allow the respondent to know that he is being measured, or that he is an object of concern to a researcher. Research has shown that this often results in changes in the behavior of the respondent (Webb, et al., 1966). As an example, the "stage" behavior that people exhibit when they realize that they are being filmed is a reaction that hardly typifies usual behavior. Another side to this problem is people's fear of being evaluated. Rosenberg has found that when people realize they are being measured, they often exhibit unusual behavior because of their concern with not appearing abnormal (1965). Rosenberg has labelled this phenomenon "evaluation apprehension", and finds that people who are high in this trait will give responses that are biased in the direction of winning the "evaluator's" favor.
- 2) The role selected by the respondent is another possible source of invalidity. This does not mean to imply that the respondent will be dishonest, but rather that he has numerous roles to fulfill, and the one selected may not be the appropriate one. For example, Orne has shown that the experimental situation contains many "demand characteristics", which elicit certain behavior from the respondent (1969). Often this behavior is to do what the experimenter asks, even if it is not the respondent's usual behavior. Support for this may be found in the recent work of Weber and Cook (1972). In a review of numerous experiments, they found that subjects adopted a "faithful" role (that is, attempting to "help" the researcher). The type of interviewing (or questionnaires) that is typically used in evaluations would appear to be particularly subject to the above two problems. It is not unlikely that the respondent would ask himself two questions: Why have I been chosen, and what do they want from me?
- 3) A final threat to the validity of questionnaires and interviews is the problem of response sets. There has long

been a critical literature on these methods that showed consistent sources of error due to the respondent's "set" to answer in a particular way. Cronbach was one of the first to summarize the area (1946). Some of the response sets he presented were acquiescence sets, where the respondent more frequently picks a positive statement than disagrees with its opposite; a set to select strong statements over moderate ones; and a set to select responses on the left side of the page more than on the right side. Cronbach also identified many idiosyncratic sets.

Attitudes and Behavior. The above discussion of several threats to the validity and generalizability of attitudinal data ignores an important question: If one does validly measure an attitude with a questionnaire or interview, just what has one measured? There is a large body of evidence that attitudes are not highly related to overt behavior. Wicker (1969) reviewed 48 studies where the respondent's verbal and overt behavior responses to objects were obtained on separate occasions. The measured attitudes were often unrelated or only slightly related to overt behaviors.

Problems With Behavioral Data. Given the above problems, it would appear that the solution would be to simply measure behavior. But in measuring behavior alone, there are also problems of validity and generalizability. For example, the data recorded is time dependent (what is obtained is a function of when the recording is done), and is spurious unless time is the variable of interest. Even more importantly, another source of invalidity is measurement as a change agent. In this case the error is not directly because the respondent reacts to being measured, but because an initial measurement causes changes in the respondent's behavior. The change itself is real, but it may be incorrectly attributed to other events (such as some aspect of the design), and subsequently incorrectly generalized to other settings that do not contain a prior measurement that caused the behavior change. The classic example of this is the well known Hawthorne studies (Roethlisberger and Dickson, 1939). The aim of the studies was to examine the effects on production of work conditions, such as temperature and illumination, at the Hawthorne plant of the Western Electric Company. The important result was that production increased no matter what condition was varied, and it did not matter how it was varied. Whether illumination was increased or decreased, production would increase. The explanation for this is the worker's knowledge that they were being studied, rather than the nature of the manipulation, resulted in the employee's behavior change.

If the people of interest are studied by observation in such a way that they cannot know that they are being measured, this problem is alleviated (e.g., Webb, *et al.*, 1966). However, any conclusions must be inferred from the behavior alone, a risky task. The obvious answer is to combine both behavioral and attitudinal measures.

Multi-Method Approach. The key to removing these problems lies in this multi-method approach. What this means is that the more ways that are

used to confirm the findings, the more certain one is of the validity of those findings. The use of questionnaires and interviews in studying the impact of designs, without the further confirmation of behavioral validation of the results (and vice versa), has left the results obtained open to many plausible, rival hypotheses. That is, there are many alternative explanations available for what was found. The evaluator cannot be sure whether the results obtained were due to the effect of the design, or come from some other real or artifactual variable(s). A general model for this multi-method approach may be found in social psychology. Campbell and Fiske (1959) have proposed a multitrait - multimethod matrix as an approach to validating experimental research. They advocate studying multiple traits, and using multiple methods to measure those traits, in order to gain validity and generalizability. If one changes "traits" to mean "behavioral constructs", the model becomes applicable to environmental assessment. Thus sound and useful information can best be obtained by always using both behavioral and attitudinal measures in an evaluation of physical settings.



# References

- Akin, O., "Contextual Fittingness of Everyday Activity Encounters," EDRA IV, 1973.
- Bechtel, R. B., "The Public Housing Environment: A few surprises," EDRA III, 1972.
- Campbell, D. T., and Fiske, D. W., "Convergent and Discriminant Validation by the Multitrait - Multimethod Matrix," Psychological Bulletin, 1959, 56, 81-105.
- Canter, D., "Office Size: An example of psychological research in architecture," The Architects' Journal, Information Library, sfB.92, Aa.3, Lancaster, April, 1968.
- Canter, D., "Royal Hospital for Sick Children: A psychological analysis," The Architects' Journal, Information Library, Cl/sfB.416, Lancaster, September, 1972.
- Cronbach, L. J., "Response Sets and Test Validity," Educational and Psychological Measurement, 1946, 6, 475-494.
- Davis, G., "Architectural Determinants of Student Satisfaction in College Residence Halls," EDRA II, 1970.
- Denton, D., et al., "Types of User Building Evaluation," EDRA IV, 1973.
- Durlak, T. J., et al., "Observation of User Patterns in Open and Traditional Plan School Environments," EDRA III, 1972.
- Esser, A. H., et al., "Territoriality of Patients on a Research Ward," Proshansky, H. M., et al., 1970.
- Hahn, R., "Behavioral Evaluation of a Juvenile Treatment Center," EDRA IV, 1973.
- Ittelson, W. H., et al., "The Environmental Psychology of the Psychiatric Ward," Proshansky, H. M., et al., 1970.
- King, L. J., Statistical Analysis in Geography, Englewood Cliffs: Prentice Hall, Inc. (Chapters 3, 4, and 5), 1969.
- Markus, et al., Building Performance, New York: John Wiley and Sons, 1972.
- Orne, M. T., "Demand Characteristics and Quasi-Controls," in R. Rosenthal and R. Rosnow, eds., Artifact in Behavioral Research, New York: Academic Press, 1969, 147-177.
- Patterson, A., "Unobtrusive Measures: Their nature and utility for architects," in Lang, et al., eds., Designing for Human Behavior, Dowden, Hutchinson, and Ross, to appear 1974.
- Rivlin, L., et al., "Age Related Differences in the Use of Space," EDRA IV, 1973.
- Roethlisberger, F. J., and Dickson, W. J., Management and the Worker, Cambridge: Harvard Press, 1939.
- Rosenberg, M. S., "When Dissonance Fails: On eliminating evaluation apprehension from attitude measurement," Journal of Personality and Social Psychology, 1965, 1, 18-42.
- Webb, E. J., et al., Unobtrusive Measures: Nonreactive research in the social sciences, Chicago: Rand McNally, 1966.
- Weber, S. J., and Cook, T. D., "Subject Effects in Laboratory Research: An examination of subject roles, demand characteristics, and valid inference," Psychological Bulletin, 1972, 77, 273-295.

## 220 methods & measures

- Wicker, A. W., "Attitudes vs. Actions: The relationship of verbal and overt behavioral responses to attitude objects," Journal of Social Issues, 1969, 25(4), 41-78.
- Wool, R., and Canter, D., "The Effect of the Meaning of Buildings on Behavior," Applied Ergonomics, 1970, 1-3.