

Chapter 20

THE EVALUATION

F. D. Becker

TABLE OF CONTENTS

I.	The Evaluation Context	232
II.	Social Sciences and Design	232
III.	Defining "Design"	232
	A. Environmental Space Management	233
IV.	The Influence of Physical Setting on Behavior	233
	A. Environment As Support: First-Order Effects	233
	B. Environment As Catalyst: Second-Order Effects	233
V.	The Environment As Social Process	234
	A. The Environment and Attention Processes	234
VI.	An Evaluation Model	235
VII.	Identifying Outcome or Effectiveness Measures	236
VIII.	"Hard" Measures	237
IX.	"Soft" Measures	237
	A. Examples of Soft Measures	238
	1. Physical Environment Checklist	238
	2. Semantic Differential	239
	3. Social and Organizational Climate Scales	240
	4. Open-ended Interview	242
	B. Examples of "Hard" Measures	242
	1. Behavioral Mapping	242
	2. Archival Records	243
X.	Simulation	244
XI.	Summary and Conclusion	245
	References	246

I. THE EVALUATION CONTEXT

A voluminous literature on evaluation and social science research exists (Strueing and Guttentag,⁴² Campbell and Stanley,¹⁰ Cook and Campbell,¹¹ Sellitz, Wrightsman, and Cook³⁶). Much of it is oriented toward the development of increasingly sophisticated methods for the collection of reliable and valid data. This literature is being supplemented by a rapidly developing literature concerned with methods for behavioral programming and postconstruction evaluation of the built environment (Preiser,²⁹ Sanoff,³⁵ Heimsath,¹⁶ Lang et al.¹⁹).

An exhaustive review of this literature, or a detailed technical description of built environment evaluation methods is beyond the scope of this chapter. Some of the basic evaluation strategies that have been applied to the study of physical settings, and hospitals in particular, will be described later, but the intent is to identify and clarify some basic conceptual issues surrounding the evaluation of physical settings. These issues include the way in which the environment is defined, some basic assumptions about environment-behavior relationships and relevant effects to measure, one's model of the planning and design process, and the relation of evaluation strategies to it. Clarifying these kinds of issues helps place in context the specific effects one attempts to measure, as well as the methods for doing so.

II. SOCIAL SCIENCES AND DESIGN

Over the last two decades an increasing number of psychologists, sociologists, urban planners, and architects have been systematically exploring the relationships between people and their physical surroundings (Sommer,³⁹ Proshansky et al.,³⁰ and Lynch²¹). The research has been "basic" in many cases, but the intent has been applied. Understanding the relationships between people and their physical surroundings was assumed to lead to better prediction and control of environmental effects on behavior. This knowledge would then be useful in planning environments that are responsive to and supportive of intended social and programmatic functions. The intent, when working at the facility level, has been organizational development, though it is seldom labeled as such (Steele⁴¹).

Changes in design practice have occurred. There is greater concern today than 15 years ago for designing buildings that are sensitive to human requirements and preferences. There is little doubt that what Carol Weiss⁴⁸ calls the "enlightenment function" of policy research, in which concepts, ideas, and information filter more or less unconsciously into policymakers' decisions, has occurred. But one can provide "input" in many ways, and to many different people involved with environmental decision making.

The orientation toward the design professions which has characterized most of the work and publications in the field of environmental psychology and sociology is beginning to broaden to include others who are involved in "environmental design". We are beginning to realize that managers and administrators are responsible for much of the day-to-day design decisions in any organization. They decide who will use what space and equipment, in what ways, when, and with whom. They decide when to renovate, or when to build a new facility. These kinds of organizational decisions are a basic form of environmental design, and it is to this audience, as much as it is to professional researchers, that this article is directed.

III. DEFINING "DESIGN"

Part of the reason managers and administrators have been overlooked as principal players in the creation of physical settings stems from a failure to distinguish between

environmental planning and design stimulated by the decision to construct a new building or do a major renovation, and the kind of planning and design required on an ongoing basis over the entire life of a building. The planning and design of new building or major renovations has received considerable attention (Preiser,²⁹ Sanoff³⁵). Ongoing environmental space management (Boutourline⁹) is often acknowledged but rarely is institutionalized in a meaningful way.

A. Environmental Space Management

Environmental space management includes decisions about major additions and renovations. It is equally concerned with decisions about the use of an existing environment or facility on an ongoing basis: decisions about which group can use the conference room; where the file cabinets should be located; how often the office landscape should be changed, by whom, and on what basis; when the corridor needs to be repainted, and in what colors; whether secretaries should be allowed into staff dining rooms, or the dining room should be separated into areas used by specific groups, and how these decisions are reached. These kinds of decisions are routinely made in organizations, usually without much attention to their programmatic ramifications: behavior and attitude of patients, staff, and visitors in the case of hospital settings, for example. Part of the problem stems from a failure to distinguish different roles the environment plays in influencing human behavior.

IV. THE INFLUENCE OF THE PHYSICAL SETTING ON BEHAVIOR

A. Environment As Support: First-Order Effects

The obvious role of the physical setting is to provide the requisite support to engage in one's job or to carry out a desired activity effectively, comfortably, and with dignity. Ranging from prosthetic devices like artificial limbs that enable persons to walk who otherwise could not, to radial corridor designs that reduce staff fatigue by decreasing distance and time traveled (Trites et al.,⁴³), the physical setting and its equipment and furnishings are generally analyzed in terms of their "first-order" or direct effect on behavior. These effects are critical. A handrail makes it possible for an infirm patient to walk independently; a narrow shower opening makes bathing of wheelchair patients difficult; the absence of appropriately sized and located storage creates the impression of chaos and poor maintenance.

Viewing the physical setting as behavioral support or barrier focuses on only one level of environmental meaning: the instrumental. It is the role of the physical setting operations researchers, human factors specialists, and other interested in "work" or "performance" take as their special province (McCormick²³).

The environment operates at several levels simultaneously, however. Our physical surroundings serve symbolic and expressive purposes as well as instrumental ones (Gibson,¹³ Sommer,³⁹ Ruesch and Kees,³³). The role of the physical setting as it serves these functions, as well as in serving the instrumental, can be analyzed in terms of "second-order" or indirect effects on behavior. With this pair of lenses attached, the focus may be on the environment as behavior *catalyst* (Becker⁶). These effects are typically masked and difficult to trace, but important because they are the link to the kinds of behavior and activity patterns that are of direct relevance to administrators, managers, and organization theorists: staff commitment, involvement, and conflict; patient trust; community support.

B. Environment as Catalyst: Second-Order Effects

The idea of environment as prosthesis, or support, is reasonably easy to understand and document. Environment as catalyst is more difficult to comprehend, and much

harder to investigate. By "catalyst" I mean that some aspect of the physical setting sets in motion a series of linked events or behavioral reactions. These reactions may be positive as well as negative. Adequate space and comfortable furniture, humidity, temperature, and lighting, in addition to directly increasing user satisfaction, may stimulate types of social interaction and communication which, in turn, provide the opportunity for staff to build patient trust, demonstrate concern, or provide timely encouragement.

The importance of the catalyst by itself may be negligible. Its value lies in its capacity to set in motion new events or processes. These events or processes, in turn, may, together and in combination, act as a catalyst for additional reactions, and so on. These second-order or "inconspicuous" effects (Steele⁴¹) can only be found if we hunt for them. They are difficult to study, in part, because identifying the "environment" that influences behavior is more problematic than it first appears.

V. THE ENVIRONMENT AS SOCIAL PROCESS

Proshansky³¹ argues that at the level of human interaction in any social setting the individual responds not to a diffusion of proximal and distal light and sound waves, shapes, and structures, objects, and spaces, but to another person, engaged in a specific activity in a specific place for a specific purpose. "Physical settings — simple or complex — evoke complex human responses in the form of feelings, attitudes, values, expectancies, and desires, and it is in this sense as well as in their known physical properties that their relationships to human experience and behavior must be understood. This is a phenomenological approach to the environment, one in which the object of study is not the environment as it is but as it is *experienced*. This is what Koffka (1935) called the 'behavioral environment.' "

When we change aspects of the physical design, we are also changing the social, administrative, and psychological environment. We are changing the total experience. A new wing of a hospital not only provides more space, but may come to mean higher status, increased management concern, or perhaps just as likely, management ignorance of appropriate or effective work procedures (Sommer³⁹). When we speak of the physical environment, we are speaking of it as it exists in this larger individual, social, and organizational context. According to Proshansky:³⁰

To speak of man's behavior in relation to the physical environment — or for that matter any kind of environment — implies that a dichotomy can be made between the person on the one hand and the environment on the other. Theoretically, however, such a distinction is untenable. There is only the total environment, of which man is simply one kind of component in relationships with other kinds of components....the physical environment, perhaps — more than any other kind, does not exist except in relation to the total environment. In speaking of the physical setting of the ward, the physical setting was not simply its physical space, its design, and the inanimate objects that occupy it. For any individual in this setting, it also included other people, their behavior, and the social context that defined what the space was for, who would use it, and what should and should not happen in it. (pp. 33-34)

Although there is only one environmental situation, there are as many surroundings as there are individuals from whose point of view the process can be examined. The kind of previous experiences we have, our current expectations, our purpose for being in a setting and our role in it, the length of time we have spent in it and expect to spend it in are, among other things, factors that insure that we continuously create the environment over time.

A. The Environment and Attention Processes

The environment, according to Weick⁴⁷ is a phenomenon tied to processes of atten-

tion. Unless something is attended to, it does not exist. Contrary to the notion that what we attend to is a function of "enduring sets," or stable characteristics that are constituted on the basis of past experience and which are carried from one situation to another, Weick argues that:

Reliance on the concept of sets imputes a reactive quality to perception and misses the point that perception creates as well as reacts to an environment. The notion that sets are relatively enduring obscures the point that perception is influenced by pragmatics, by the particular here and now in which it originates. Perception is a more complex and general phenomenon than is attention. Thus a study of attention would appear to be a more logical and manageable starting point for an inquiry into the ways in which organizational environments are created. When we study attention, it is not important to discover immediately all prior experiences that an actor has had. Instead, the immediate question concerns what is happening in the actor's present situation that controls the nature of the attention he directs to his past experience. (p. 39)

Meaning, according to Weick, is the kind of attention directed to the past; that is, attention to actions we have already taken. One's current interest and concerns influence the events, processes, and behavior we choose to focus our attention on. Any action performed in the past may be subject to changing interpretations over time as a function of changing interests or concerns. And since any action carries "surplus meaning" (our attention is directed only to a portion of the total possible behaviors we have engaged in), portions of the stream of experience which went unnoticed at the moment of the reflective glance can still be noticed on some future occasion.

I have borrowed Weick's conception of organizing processes because it seems to me especially useful in shifting the focus of attention from the physical setting as a *product*, which is viewed as something relatively permanent and unresponsive and largely existing outside of everyday ongoing social processes, to the physical setting as a dynamic ongoing social *process*. As such, the physical setting can no longer be accepted as a "given." Its meaning, and the role it plays, becomes as open to question as does the way in which it develops. Given this way of conceptualizing the environment, the next question is how we conceptualize the planning process and its relation to evaluation.

VI. AN EVALUATION MODEL

Ideally evaluation should support the planning process. To do so, models of each should be congruent. For complex settings like hospitals a cybernetics model of the planning process (Bauer³) seems especially useful. The word "cybernetics" comes from the Greek "helmsman" and reflects the constant adjustments of course that characterize the way in which a sensitive helmsman controls the ship. When the wind changes direction, the helmsman changes the task or misses the destination point. Winds are unpredictable and change all the time, as are the forces influencing behavior in any organization. The essence of this model is the assumption that one can achieve at best only some *approximation* of what one aims at, based on an initial course of action. As Bauer notes, salvation lies in rapid detection of error and adjustment to correct for that error. The value of this planning approach is that it removes the stigma of not attaining a goal in an anticipated manner:

This cybernetic model also makes it natural to think of the second-order consequences of one's actions. "Error" is not a rare phenomenon that occurs because of bad planning and inept control. It is the natural and inevitable feature of all purposive action. Ironically, this technological modern view of planning and control gives respectability to the once derogated British doctrine of "muddling through." The British doctrine of muddling through might be described as deciding generally where one wants to go, taking a step in that direction, reassessing the consequences, reassessing one's goals and methods, taking another step, and so on. (p.25)

This model of the planning process suggests that evaluation should be continual, relatively rapid, and done in the context of administrative mechanisms that facilitate change. A major goal becomes rapid identification of problem areas, rather than the development of knowledge about basic relationships among variables. Evaluation of this sort does not necessarily conform to canons of scientific research, and need not do so to be useful to the organization. Evaluation in this context is "sense-making" (McClintock²²).

The primary audience is the organization itself, not some larger professional scientific audience. The results of these kinds of investigations may, however, guide and inform more rigorous scientific research that has as a primary aim the establishment of basic relationships among variables. This kind of research is generally characterized by more sophisticated research designs, larger sample sizes, sophisticated selection, and statistical techniques, and so on.

This kind of research is expensive because it occurs in extended time frames, and requires a large number of persons with a high level of specialized research experience and skill. For this reason, it is infrequently done as a matter of course in organizations. Examples of both types of research orientation will be presented below. The main point here is that the less expensive, more regularly occurring form of evaluation is a legitimate and appropriate evaluation strategy within a cybernetic model of the planning process. Although as decisions become more costly, in either human or dollar terms, we would like to make them with greater degrees of certainty about their effects, the fact of the matter is that even the more rigorous and "scientific" research rarely provides unequivocal answers to complex questions about the often subtle relationships between the environment and behavior. Through collaboration, cooperation, and mutual assistance, both of these forms of design research can occur. They complement, rather than substitute, for each other.

Having set the context for the evaluation of hospital environment by briefly describing a framework for defining the environment, and a model of the planning and evaluation process, we can now turn to a consideration of *what* should be studied, and methods for doing so.

Our purpose here is to discuss, first, some of the issues surrounding the kinds of measures one chooses to use, and then to describe collection techniques that have been specifically developed and applied to the evaluation of physical settings, particularly hospital environments.

VII. IDENTIFYING OUTCOME OR EFFECTIVENESS MEASURES

We tend to overlook the fact that an "outcome" is any observable datum one chooses to focus on from all those potentially available in a situation. The datum may be observed spontaneously, as the behavior occurs; it can be elicited for specific purposes, as in the case of testing and experimental situations; or it can be inferred from other behavior or signs of behavior, as in the case of archival records (e.g., absenteeism, turnover) or physical evidence of use or abuse of the environment reflected in measures of changes in environmental conditions over time (e.g., damage, materials wear, altered arrangements, modifications and additions). These data may be observed and or measured in the specific setting of interest (e.g., office or classroom) or in some setting theoretically or empirically related to it (e.g., children's or employee's home environment).

In a hospital setting these "outcomes" can include, but are not limited to: attitudes about health, doctors, nurses, or one's self; social behaviors among staff such as cooperation and conflict, social support, independence, and personal initiative; administrative and managerial behaviors including amount of time spent planning and organ-

izing activities as a percentage of total time helping patients or doing relevant paper work; lists of number and type of rules and how often they are broken; average time spent in recovery; duration before onset of postsurgery ambulation; staff turnover and absence rates; perception of spaciousness or comfort or detail under different lighting conditions. The list is endless.

Selection is a function of values. When multiple outcomes are observed, as is increasingly the case, these have been selected from an enormous pool of potential outcomes. We focus on patients' or physicians' subjective reactions to lighting or noise because we believe there is a relation between these perceptions and performance, stress, or health. Different values direct our attention to different outcome measures. Since values are closely linked to role requirements and professional identity and training, it is unsurprising that physicians, administrators, nursing staff, and housekeeping staff, not to say patients and their visitors, focus on different aspects of the total situation and value some outcomes more than others.

VIII. "HARD" MEASURES

Some measures are also more "culturally powerful" than others (Becker⁵). Administrators find figures about absenteeism and turnover more persuasive than graphs of staff's attitudes toward work; measures of ambulation or recovery rate are typically more valued than measures of patients' perceived comfort. The more valued measures are generally "hard" rather than "soft"; that is, they typically refer to something that can be directly observed in some fashion, and which is not dependent on the respondent providing the information, as is the case with subjective responses.

The problem, of course, is identifying relevant behaviors and devising ways to measure them. A number of different "hard" measures have been used. Some of these include: nurse travel time and travel distance (Trites et al.,⁴⁴ Lippert²⁰); individual interaction patterns between staff and staff and patients (Becker and Poe⁶); turnover and absenteeism rates (Trites et al.⁴⁴); time to complete different tasks and intensity of equipment and space use (Rawlinson³²); average cost and length of stay (Beckman⁸).

The kinds of measures selected are related to different focuses of concern: patients are more likely to be concerned about nurse-patient interaction than intensity of use; administrators are more likely to be concerned about intensity of use and absenteeism than the quality of interaction between nurses and housekeeper; nursing staff is more likely to be concerned about travel distance and time than absenteeism. No single measure is inherently more appropriate or valid than any other. Ideally, a comprehensive evaluation would include measures of interest and relevance to the various participants in the setting: administrators, nursing, housekeeping, and maintenance staff; patients and visitors; and so on. To be avoided is the situation where any single set of criteria become the only standards for evaluating the entire performance of the setting. "Hard" measures are powerful and persuasive, but they do not tell the whole story. There is also a need for "soft" or subjective responses; that is, self report.

IX. "SOFT" MEASURES

Although generally less powerful than direct observations of behavior, "soft" or subjective responses are valuable in evaluating environments for several reasons. Given the framework for defining the environment developed earlier, with its emphasis on attention processes, there is a need to identify which aspects of the setting people attend to. One way of doing this, and often the most economical in terms of time and energy, is to ask persons what is important, what they notice. Observations reveal behavior patterns, but tell us little about the specific aspect of the environment that may be

influencing the behavior unless comparisons are made under carefully controlled conditions where particular aspects of the environment are systematically varied. This is certainly possible, but often requires a research commitment most organizations cannot support.

Eliciting user's subjective responses to the environment is also useful in helping to identify not only what they are attending to, but what the nature of the problem is with some aspect of the setting (Becker⁵). Such information is very useful in suggesting directions for change in the design or management of the space. Some reactions of interest are also extremely difficult or impossible to observe directly: apprehension, confidence, depression, satisfaction, feeling of competence, or control. These kinds of subjective reactions are often explicit objectives of environmental design decisions, and can be measured directly by asking people to report on their feelings. "Soft" measures are also useful in identifying different individual's or group's objectives, and the meaning they attach to certain characteristics of the environment. Direct observation may identify environmental dysfunctions, but they are less helpful in defining the importance or meaning of the dysfunction; that is, in determining priorities among environmental design elements that might be changed.

As Webb⁴⁵ noted several years ago no single method is perfect. Each has its own source of error. The solution is what they called "triangulation," or the use of multiple methods, each with its own but different source of error. The different methods, in combination, provide a picture of the phenomena that any single method can never do. Direct observations are useful when behavior responses and patterns are outside of awareness, when individuals are unable to make verbal reports (e.g., infants, retarded), and when there is some reason to assume individuals may want to distort their verbal responses (e.g., time-activity diaries underrepresent sexual and conflict behaviors). Because it is often easier to develop simple questionnaires than to operationalize procedures for observing subtle and complex behaviors, train observers, and analyze thousands of behavioral episodes, social science research has relied on subjective responses. There is a need to redress the balance without underestimating the value of "soft" measures for eliciting useful information.

A. Examples of Soft Measures

The following is a selected sample of instruments that have been used to measure subjective responses to the built environment. The intent is to provide a sense of the strengths of different types of instruments, not a technical description.

1. Physical Environment Checklist

Sommer⁴⁰ has used this type of instrument in a wide variety of settings, and Becker and Poe⁶ adopted it to the study of design changes in a small general hospital. It can vary greatly in detail and length, but typically it lists areas of the physical setting (e.g., lounge, patient room) and specific aspects of the setting (e.g., lighting, noise, temperature, ventilation), as well as specific behaviors of interest (e.g., cooperation, friendliness of staff (see Table 1). Respondents indicate their level of satisfaction for each item on the checklist, and also what it is about each item that is especially good or bad (e.g., the unpredictability, not just level of noise). This information directs attention to the specific aspect of the environment that is important to the individual, and guards against the possibility that characteristics individuals especially value are inadvertently eliminated when correcting other problems. Respondents can also be asked to indicate the importance to them of each area or characteristic of the environment. This information is valuable in establishing priorities for decision making. Scoring involves tabulation of percentages in each rating category, and listing of specific prob-

Table 1
PHYSICAL ENVIRONMENT CHECKLIST*

Specific Areas	Needs			What is especially good or What needs improvement?
	Excellent	Satisfactory	Improvement	
A. Corridor				
B. Visitors waiting area				
C. Patient room				
D. Porch				
E. Nursing Station				
F. Shower room				
G. Area outside elevators				
H. Bathrooms in patient rooms				

Note: Ratings of importance for each area, and more specific elements within areas (e.g. color of walls, artificial light, equipment noise) can easily be added to this form.

* From Becker, F. D. et al. *User Participation, Personalization, and Environmental Meaning*. Ithaca, NY: Program in Urban and Regional Studies, Cornell University, 1978.

lems and advantages of individual items. It is very easy to develop the items, and it is quick and easy to complete and score. For these reasons, it is a very useful instrument.

2. Semantic Differential

Like the Physical Environment Checklist, the semantic differential has been widely used in evaluations of physical settings because of its ease of construction, administration, and scoring (Sanoff³⁵). While responses can be factor analysed to identify major variables and dimensions, it is often used in a fashion similar to the Checklist. Bipolar adjective scales are developed that represent concepts considered as relevant characteristics or objectives of whatever is being evaluated. For example, a day room may be evaluated on "cheerful — depressing," "comfortable — uncomfortable," "orderly — chaotic" scales.

Typically, each scale has seven points along it (see Table 2), and more than one scale is used to tap a few basic dimensions. For example, in a study of the impact of different lighting arrangements on spatial impression, Flynn et al.¹² used three dimensions, each composed of several different bipolar adjective scales (i.e., Evaluation included friendly-hostile, pleasant-unpleasant, like-dislike, relaxed-tense; Perceptual Clarity included clear-hazy, distinct-vague, and bright-dim; and Spaciousness included large-small, short-long, spacious-cramped) to measure the effects of the lighting arrangements. Differences between groups, (e.g., nurses and patients) perceptions of the same space, or the same group's perceptions of different physical areas can easily be identified by comparing the arithmetic mean of different group's responses to different scales. Flynn et al.¹² found, for example, that limited wall lighting significantly affected the impression of spaciousness and perceptual clarity.

The semantic differential is most useful when one knows the specific aspect of the physical setting to which the individual is responding. Unfortunately, this is often not the case. Two individuals may have the same numerical score on a scale, but be responding to entirely different aspects of the setting. Unless used in conjunction with other instruments, or as part of a controlled study in which the nature of the stimulus is varied systematically so that any observed differences in scores can be attributed to differences in the stimulus, the semantic differential is primarily helpful in identifying basic problems or areas of excellence. It is useful to know, however, not only whether a room or environmental characteristic is pleasant or unpleasant, but also whether this

Table 2
SEMANTIC DIFFERENTIAL
EVALUATION OF DIFFERENT
LIGHTING ARRANGEMENTS*

Evaluative dimension	
Friendly _____	Hostile _____
Pleasant _____	Unpleasant _____
Like _____	Dislike _____
Harmony _____	Discord _____
Satisfying _____	Frustrating _____
Beautiful _____	Ugly _____
Sociable _____	Unsociable _____
Relaxed _____	Tense _____
Interesting _____	Monotonous _____
Perceptual Clarity dimension	
Clear _____	Hazy _____
Bright _____	Dim _____
Faces Clear _____	Faces Obscure _____
Distinct _____	Vague _____
Focused _____	Unfocused _____
Radiant _____	Dull _____
Spaciousness	
Large _____	Small _____
Long _____	Short _____
Spacious _____	Cramped _____

* From Flynn, J. et al., *J. IES*, 87, 1973. With permission.

is a dimension that is important to the individual. Again, such information helps set priorities and future courses of action.

3. Social and Organizational Climate Scales

While both the semantic differential and environmental checklist have been widely applied to *direct* evaluations or ratings of the physical setting, social climate scales (Moos²⁷) and organizational climate scales (Becker⁵) elicit information indirectly about the effects of the environment. Differences in scores between different groups evaluating the same setting (or the same setting before and after some environmental modification) are attributed to differences in the physical setting, when other factors such as type of population or health status of respondents have been controlled experimentally or statistically.

For example, Moos²⁷ developed a Ward Atmosphere Scale to measure the social climates of psychiatric treatment programs as perceived by patients and staff. Drawing on trained observers who observed three different wards for several weeks, reviews of popular and professional books, and interviews with patients and staff, an initial pool of more than 500 items was developed that described important dimensions of the treatment setting. In the case of the psychiatric facility, some of these dimensions included *Involvement*, or the extent to which patients are active and energetic in the day-to-day social functioning of the ward; *Support*, or how helpful and supportive patients are toward other patients, or doctors are toward patients; *Autonomy* assesses how self-sufficient and independent patients are encouraged to be in their personal affairs and

Table 3
A PARTIAL LIST OF DIMENSIONS AND
STATEMENTS FROM THE WARD ATMOSPHERE
SCALE

Involvement Subscale

- The patients are proud of this ward
- Nobody ever volunteers around here

Support Subscale

- Staff go out of their way to help patients
- Doctors have very little time to encourage patients

Order and Organization Subscale

- This is a very well organized ward
- The day room is often messy

Program Clarity Subscale

- People are always changing their minds here
- Ward rules are clearly understood by patients

Anger and Aggression Subscale

- Patients often gripe
- It's hard to get people to argue around here

in their relationships with staff, and the openness of staff to patients' suggestions. Other dimensions include Spontaneity, Practical Orientation, Personal Problem Orientation, Order and Organization, Program Clarity, and Staff Control.

From the initial pool of statements, and following extensive testing to cull out overlapping, confusing, or otherwise inappropriate items, a final long and short version of the Ward Atmosphere Scale was developed. Respondents are given a form with the statements (see Table 3) and asked to decide which statements are true of their ward and which are not. From these responses a profile of different groups' perceptions of the same environment is developed (see Figure 1).

In a variation of the Moos approach, Becker and Poe⁶ developed a brief questionnaire consisting of statements to which respondents indicated their degree of agreement or disagreement along a seven-point scale, and administered it before and after a renovation of a nursing unit to patients, nursing and housekeeping staff, and visitors (see Table 4). Statements were generated from informal observations and conversations on the nursing unit. These interviews indicated, for example, that for some users the hospital was a boring, impersonal, and depressing place. Consequently, items were formed to tap these and other feelings that were felt might be altered by the renovation. Five general categories of attitudes were selected: (1) mood and morale; (2) perceived quality of health care; (3) attitude toward the hospital administration; (4) attitude toward the nursing staff, and (5) environmental awareness. Because of limited time, the categories were not intended to represent statistically determined scales (Moos²⁷), but rather to provide cross-checks on responses to individual items. Pre- and posttest means were calculated for the same groups, and for different groups, and comparisons were made with two nearly identical nursing units that served as controls. They found, for example, that the physical renovation had a positive effect on general mood and morale when the responses of all users were combined, but that staff were affected most positively. Patients also showed an increase in the "perceived quality of health care," but staff showed an even greater increase.

The social climate scales are more difficult to develop than either of the above instruments, since the wording of the statements is crucial. The development of reliable and valid scales is a project in itself. Unless one uses scales already developed, this approach to evaluation is beyond the reach of most organizations. Such scales need to be used in conjunction with other methods that identify specific aspects of the environment individuals are responding to, and the meaning it has for them.

Table 4
A PARTIAL LIST OF DIMENSIONS AND STATEMENTS FROM AN
ORGANIZATIONAL CLIMATE QUESTIONNAIRE IN A GENERAL
HOSPITAL

- Mood/Morale
 - The hospital environment makes me apprehensive
 - The noise around here makes me tense and nervous
- Quality of Health Care
 - This hospital seems a little behind the times
 - I have confidence in the quality of medical care provided here
- Organizational Issues
 - The hospital seems impersonal and bureaucratic
 - This hospital is well organized
- Nursing Staff
 - The nursing staff has little patience
 - The nursing staff is helpful and cooperative

4. Open-ended Interview

Sommer⁴⁰ reports a study by Dudgeon and Davidson (1965) of reactions of patients to their stay in hospital in which open-ended interviews are used. One hundred and fifty patients were interviewed and simply asked to comment about the physical condition of the hospital, including the size of the wards, lighting, ventilation, the bathrooms, and so on. The questions were later coded into categories by the interviewer. Particularly in an exploratory study, where the main intent is to get a feel for the kinds of problems and concerns of different users, these kinds of open ended questions are useful. This kind of instrument is very useful and economical to use in a small-scale survey dealing with a specific building and aimed at a finite population. Unlike the social climate type of questionnaire, it does not require a heavy commitment of research funds or considerable technical research experience.

These few instruments by no means exhaust the range of methods that can be used to evaluate subjective responses to the environment. Several sources provide a more complete description of common techniques for measuring subjective responses to the environment (e.g., Ostrander,²⁸ Sanoff,³⁵ Lang et al.,¹⁹ Michelson²⁵). We turn now to a similarly selected list and discussion of techniques for collecting "hard" data.

B. Examples of "Hard" Measures

By "hard" measures I mean ones that do not depend on an individual voluntarily and consciously providing information about their own behavior, attitudes, or feelings. There are two basic forms of "hard" data: direct observation of behavior (recorded by human observers or instruments like cameras, tape recorders, etc.), and the use of archival data, or data that has been collected routinely as part of an organization's normal operating procedure.

1. Behavioral Mapping

Behavioral mapping (Ittelson et al.,¹⁸) is a method of systematic observation in which an observer moving through a space in a predetermined sequence notes who (e.g., patients, staff, visitors) is doing what (e.g., talking, charts, patient consultation, reading) where (bed in patient room, couch in patient lounge, table in nursing station), when (morning, evening, night, day of week). It can also include the size of interacting groups. It does not record behavior over time in a fixed location. Becker & Poe,⁶ for example, obtained architectural plans of three target wings in the general hospital where they did a small renovation of one nursing unit. From these plans scale maps

of the wards were drawn. At prescribed intervals, an experimenter made tours of the three wards using the maps to record the position, group status, and activity of each person encountered. Each of the wards was toured three times on any given visit to the hospital, with the touring order randomized within blocks of three tours. A simple coding system was used for user group classification (e.g., N for nurses, P for patients, M for male, F for female) and for type of activity (e.g., T for talking, R for reading). They found increased use of the solarium, the social area that had received the most extensive design changes, particularly for patients during the day. The changes had little effect on staff members, who rarely used this space at any time. This information is useful since one objective was to make patients recovery less boring. Changes in the use of the solarium created improved opportunities for stimulation and socialization.

In a larger, landmark study evaluating the effects of different types of corridor arrangements on the activities and subjective feelings of nursing personnel, Trites et al.⁴³ used different observational methods to record nurse travel time and distance. Using 14 senior nursing students who had participated in an intensive 2 week training period in work sampling techniques, data on the type and location of nursing staff activities was collected by the student observers (see Trites, et al.,⁴⁴ for a thorough description of the rather complex procedures used).

The data was collected by counting the number of trips to patient rooms in units of each design on eight or more day shifts. The distance figure was derived from a combination of average distances to patients rooms and average numbers of such trips per staff member under the assumption that all trips to patient rooms started from the nurses station. Using these derived figures (which are open to question, since nurses often do a "tour" (Lippert²⁰) in which several patients are visited in sequence without returning to the nurses station) they found that the average distance traveled on the radial units was 0.61 miles. Using this figure they calculated that to limit the distance traveled on single and double loaded corridors to this same figure would require an additional 1.75 people on the double corridor unit and 4.6 people on the single corridor unit.

These data are observational but lose some of the power of other types of observational data because comparisons are made on the basis of indexes that depend on certain assumptions holding true (e.g., nurses visiting patients' rooms start all trips from the nurse's station). They also depend on more sophisticated techniques for collecting data than the typical type of behavioral mapping reported above. This means they are less likely to occur than more simple evaluations done on an ongoing basis by an existing hospital staff or outside consultant brought in for a short period.

2. Archival Records

Most organizations collect an enormous amount of data for routine reporting purposes. Rarely are these data used as an integral part of an evaluation program, but they can be. In the study already noted by Trites et al.⁴³ another measure of the effect of corridor designs (radial, single, and double loaded) on nurses activities was the use of records on staff absenteeism and accidents. They found significant differences among the nursing-unit designs. There were relatively few absences on radial units, next fewest on double corridor units, and the most on single corridor units. More accidents occurred on single corridor units compared with radial and double corridor units.

Other types of archival data includes formal complaints and requests for transfer, and length of patient stay. These data are economical to collect since they are routinely recorded as part of the hospital's operating procedures. They are "hard" since they are not dependent on self-reports of feelings and attitudes. They are not particularly helpful, however, in pinpointing the specific aspect of the setting that may be causing

a problem. For this reason, like several of the methods mentioned above, they must be used in conjunction with other kinds of information.

X. SIMULATION

In many instances one would like to know before a substantial investment of energy and funds whether or not the planned changes will have the desired effects. Environmental simulations, or the presentation of essential components of the planned change in some form, to which the actual or prospective users of the space or equipment can react, are one possibility. Several simulation techniques have been used to study response to planned environmental change (McKechnie²⁴). Like any other method, each has its own set of strengths and weaknesses.

Color or black and white slides to which individuals respond on structured questionnaires using either semantic differential or an activity checklist (Sims³⁷) can be used to elicit reactions to proposed changes in the built environment. Using a photomontage (Sirlin³⁸) process in which an existing photograph has superimposed on it the image of the proposed change is another way of visually presenting such changes. Detailed drawings of the proposed changes can also be made directly on the slide so that the changes are viewed within the specific context in which they will in fact exist (Hascup¹⁵). Sims³⁷ has found that color renderings of proposed changes are the most effective in eliciting responses closest to actually viewing the environment directly. Models can also be used, and viewers' reactions aided by manually using a special optical viewing device that the viewer runs along the bottom of the model while looking into an eyepiece. The viewer sees the model as though he or she were walking through it, as opposed to viewing it from a bird's eye perspective. To provide a perceptually dynamic view of the planned environment (Appleyard and Craik²) that appears extremely lifelike, computer graphics (Greenberg¹⁴) in which a computer stores information on the location of buildings, rooms, corridors, and so on are also possible. Using this stored information, the computer can generate a cartoon-like movie that presents a planned environment as through one were walking through it, or seeing it from different angles and approaches. These systems, at present, are expensive, time-consuming to develop, and require an enormous amount of sophisticated computer programming.

To these various forms of presenting the environment, individuals can be asked to indicate their aesthetic response, activities they believe they might engage in, paths they might take through the environment, and so on (Sims,³⁷ Appleyard and Craik²).

Some studies have found subjective responses quite close to those obtained from responses to the actual situation (Howard et al.¹⁷). The absence of a time dimension, of other people acting in the environment simultaneously, and a specific social context which directs attention to some aspects rather than others makes these kinds of simulations less useful when viewers are unfamiliar with the proposed environment, or interest is more in behavioral response (e.g., performance, activity patterns). Simulations provide a useful insight into possible responses but their value can be enhanced by using them in conjunction with other types of preconstruction evaluation.

Full-scale mock-up of patient rooms may be done at relatively low cost (Wehrer⁴⁶), or nurses or patients might travel to nearby locations where hospitals with some of the proposed features can be observed, staff and patients interviewed and so on. Too often, even when a full-scale mock-up is constructed, it is not effectively evaluated. In one case with which I was involved, a patient's room had been built, but it was in a distant part of the hospital facility, and neither nurses or patients were taken to it, (or better, used it) or were asked to evaluate it on a number of dimensions of interest. Essentially, it was just there, and a few nurses were asked to comment about it by the architects.

Although not always feasible, another useful strategy is to build and design incrementally (Alexander et al.¹) so that users' responses to part of a fully operative environment can be used as a basis for design decisions in subsequent construction phases. Inexpensive simulations placed in existing environments are still another possibility. One might hang inexpensive panels with proposed color schemes in existing hallways so users can experience and respond to them as they will function in the new environment. Seeing a paint chip, or even a color rendering of a room, is no substitute for experiencing the proposed change at life size in the actual context.

XI. SUMMARY AND CONCLUSION

The preceding discussion suggests a number of implications for evaluating hospital environments. Assuming that the environment is not a "given," but is created by participant through those aspects they choose to attend to:

1. Researchers need to identify those aspects of the setting individuals attend to. It is useful to ask participants what it is about the environment that is causing a problem, and ask them to describe how they cope with it.

While people have little trouble indicating their level of satisfaction with specified environmental elements, they may be less aware of how aspects of the environment are affecting their behavior. For this reason:

2. Researchers need to increase the use of systematic observational procedures that are not dependent on self-reports. If the goal is to attribute observed behaviors to particular aspects of the environment, careful comparisons of situations with and without these elements will have to be made. Small scale ministudies in which changes are done incrementally is a useful strategy, as are before-after comparisons in the same situation. The process of change is part of the change, and needs to be given as careful consideration as the nature of the changes themselves.

The environment operates at different levels; that is, it may stimulate direct first-order effects as well as more indirect second-order effects. Rarely do evaluation studies investigate the more subtle second-order effects, and yet these are often of most interest to hospital administrators and nursing staff. For example, while increased travel distance and time was associated with nonradial corridor design (Trites, et al.⁴⁰), and its importance discussed in terms of increased staffing loads, no effort was made to assess systematically the impact of travel distance and time on nurse-patient interaction, or the quality of care received. Identifying these subtle, inconspicuous second-order effects is difficult and often time-consuming, but:

3. Increased effort should be devoted to the investigation of second-order consequences of environmental characteristics. These efforts, when done systematically, are more likely to require a large commitment to research, and so may be less typical of the research that can be done by in-house hospital staff. Yet, because these effects are subtle, they may be more noticeable to permanent observers than transient research teams. Effort should be made to identify the ways in which individuals cope with and adapt to different environmental configurations.
- 3a. Second-order effects may not occur in the target setting. The effect of new furniture and lighting in patient rooms may be noticed, for example, in changes in patient behavior in the lounges and day rooms. Therefore, linked settings, or systems of settings, should be considered as research sites, not only single settings.

Because research is expensive when done as a totally separate operating activity, every effort should be made to use existing resources. Therefore,

4. Increased use should be made of archival data; that is, hospital records about staff absenteeism, accidents, requests for transfer; patient length of stay; ambulation after surgery, and so on. The form in which the data is collected and stored should be considered in light of potential in-house research applications. Existing personnel resources should also be used imaginatively. For example, in the Trites et al.⁴⁴ study, nursing students collected the data as part of their educational requirements. Hospitals that sponsor education programs of any sort should be alert to student's potential as researchers.

Because the environment is a social process, influenced by the nature of the social system which it both helps to define and is defined by,

5. Evaluation of hospital environments should also take into consideration the nature of the social system. Relations among different user groups, the system of authority and control characterizing the administration, and so on, should be assessed as well as the environment itself. Differences in this social system often account for similar or identical environments having very different observed effects.

Given the cybernetic model of planning, in which mistakes are assumed to occur regardless of the nature of planning involved, hospitals should seriously consider ways in which they can develop environmental monitoring systems. Their purpose is to identify environmental dysfunction and negative consequences at early stages, before they become stable patterns treated as inevitable and unchangeable. Some of the problems may be alleviated by simple environmental changes, as they were in our own study (Becker⁷). In other cases, changes in the use and management of the space may be more appropriate or feasible. And in still other cases, the system will identify priorities for future, larger-scale renovations. Some of the most simple and obviously environmentally related problems are never rectified because there is no one to coordinate a planning effort. The environmental monitoring system provides such a coordinating mechanism.

I have tried to make clear that evaluation can occur at many different levels. The more the evaluation moves toward the goal of "science", or the development of basic propositions, the more expensive and time-consuming the research becomes. As it moves more toward the "sense-making" end of the scale, where the objective is more to understand one's own situation and do everything possible to improve it, the less expensive and time-consuming is the research. Both approaches are valid. Too often, the value of the simpler, smaller scale research is overlooked.

REFERENCES

1. Alexander, C., et al., *The Oregon Experiment*, N.Y., Oxford Univ. Press, 1975.
2. Appleyard, D. and Crails, K., The Berkeley environmental simulation project: its use in environmental impact assessment, in *Environmental Impact Assessment: Guidelines and Commentary*, Dickert, T. G., and Domeny, K. R., Eds., Berkeley: Univ. Extension, U.C. Berkeley, 1974.
3. Bauer, R., *Second-order consequences*, MIT Press, Cambridge, Mass. 1966.
4. Beck, W., *Operating room lighting*, American Sterilizer Company, Erie, Pa., 1966.

5. Becker, F. D., Building by design, *Hum. Ecol. Forum*, 9, 2, 3—7, 1978.
6. Becker, F. D. and Poe D., The effect of user-generated design modifications in a general hospital, *J. Nonverbal Behav.*, 4, 4, 1980.
7. Becker, F. D., *Workspace: Creating Environments in Organizations*, Praeger, N.Y., 1981.
8. Beckman, R., Getting up and getting out, *Prog. Archit.*, Dec. 64, 1974.
9. Boutourline, S., The concept of environmental management, in *Environmental Psychology*, Proshansky, H., Ittelson, W., and L. Rivlin, Eds., Holt, Reinhart, and Winston, N.Y., 1970, 496.
10. Campbell, D. and Stanley, J., *Experimental and Quasi-Experimental Designs for Research*, Chicago, Rand-McNally, 1966.
11. Cook, T. and D. Campbell, *Quasi-Experimentation Design and Analysis Issues for Field Settings*, Rand-McNally, Chicago, 1979.
12. Flynn, J., Spencer, T., Martynuk, O., and Hendrick, C., Interim study of procedures for investigating the effect of light on impression and behavior, *J. IES*, Oct., 87, 1973.
13. Gibson, J. J., *The Perception of the Visual World*, Houghton Mifflin, Boston, 1950.
14. Greenberg, D., *Computer graphics in architecture*, *Scient. Am.*, 1974, 230, 98, 1974.
15. Hascup, G., Personal communication, November, 1973.
16. Heimsath, C., *Behavioral Architecture*, McGraw-Hill, N.Y., 1977.
17. Howard, R., Mlynarski, G., and Sauer, G., A comparative analysis of effective responses to real and represented environments, Proceedings of EDRA 3 conference, UCLA, January 1972.
18. Ittelson, W., Rivlin, L., and Proshansky, H., The use of behavioral maps in environmental psychology, in *Environmental Psychology*, Proshansky, H., Ittelson, W., and Rivlin, L., Eds., Holt, Rinehart, & Winston, N.Y., 1970, 658.
19. Lang, J., Burnette, C., Moleski, W., and Vachon, D., *Designing for Human Behavior*, Dowden, Hutchinson, & Ross, Stroudsburg, Pa., 1974.
20. Lippert, S., Travel in nursing units, *Human Factors*, 13, 3; 269, 1971.
21. Lynch, K., *Image of the City*, MIT Press, Cambridge, Mass. 1960.
22. McClintock, C., Program evaluation as sense-making: an approach for the 1980's. Paper presented to Human Services Studies Department, College of Human Ecology, Cornell University, Ithaca, New York, 1979.
23. McCormick, E., *Human Factors in Engineering and Design*, McGraw-Hill, N.Y., 1976.
24. McKechnie, G., Simulation Techniques in Environmental Psychology, in, *Perspectives on Environment and Behavior*, Stokols, D. Ed., Plenum Press, N.Y., 1977.
25. Michelson, W., Ed., *Behavioral Research Methods in Environmental Design*, Dawden, Hutchinson, & Ross, Stroudsburg, Pa., 1975.
26. Moos, R., *The Social Climate Scales: An Overview*, Consulting Psychologists, Inc., Palo Alto, Calif., 1974.
27. Moos, R., *Evaluating Treatment Environments: A Social Ecology Approach*, NY: Wiley, N.Y., 1974.
28. Ostrander, E., *Methods Notebook: A Brief Description of User Reaction Research Instruments*, Unpublished manuscript, Cornell University, Ithaca, N.Y., 1972.
29. Preiser, W., *Facility Programming*, Dawden, Hutchinson, & Ross, Stroudsburg, Pa., 1978.
30. Proshansky, H., Ittelson, W., and Rivlin, Eds., *Environmental Psychology: Man and His Physical Setting*, Holt, Rinehart, and Winston, N.Y., 1970.
31. Proshansky, H., Ittelson, W., and Rivlin, L., The influence of the physical environment on behavior: some basic assumptions, in *Environmental Psychology*, Proshansky, H., Ittelson, W., and Rivlin, L., Eds., Holt, Rinehart, & Winston, N.Y., 1970.
32. Rawlinson, C., Space utilization in hospitals, *J. Archit. Res.*, 6, 3, 1978.
33. Ruesch, J. and Kees, W., *Nonverbal Communication*, U.C. Press, Berkeley, Calif., 1964.
34. Sims, W., et al., A methodology for assessing the effectiveness of proposed sign regulations, Proceedings of EDRA 9 conference, Buffalo, NY, April 1979.
35. Sanoff, H., *Methods of Architectural Programming*, Dawden, Hutchinson, & Ross, Stroudsburg, Pa., 1977.
36. Sellitz, C., Wrightsman, L., and Cook, S., *Research Methods in Social Relations*, 3rd. Edition, Holt, Rinehart and Winston, N.Y., 1976.
37. Sims, W., Iconic Simulations: An Evaluation of their Effectiveness as Techniques for Simulating Environmental Experience Along Cognitive, Affective, and Behavioral Dimensions, unpublished Ph.D. dissertation, Dept. of Urban Studies and Planning, MIT, 1973.
38. Sirlin, J., Photomontage broadens the spectrum of design, *Progr. Archit.*, Jan., 56, 81.
39. Sommer, R., *Personal Space: The Behavioral Basis of Design*, Prentice-Hall, Englewood Cliffs, N.J., 1969.
40. Sommer, R., *Design Awareness*, Rinehart Press, San Francisco, 1972.
41. Steele, F., *Physical Settings and Organization Development*, Addison-Wesley, Reading, Mass., 1973.
42. Struening, E. and Guttentag, M., *Handbook of Evaluation Research*, Sage Publications, Beverly Hills, 1975.

43. Trites, D., Galbraith, F., Leckwart, J., and Sturdavant, M., Radial nursing units prove best in controlled study, *Mod. Hosp.*, 112, 4; 94—99, 1969.
44. Trites, D., Galbraith, F., Sturdavant, M., and Leckwart, J., Influence of nursing-unit design on the activities and subjective feelings of nursing personnel, *Environ. Behav.*, 1970, 2, 3, 303, 1970.
45. Webb, E., Campbell, D., Schwartz, R., and Sechrest, L., *Unobtrusive Measures: Nonreactive Research on the Social Sciences*, Rand-McNally, Chicago, 1966.
46. Wehrer, J., A design research strategy utilizing simulated health care procedures in an environmental mock-up unit, *Proceedings of EDRA 3*, UCLA, 1972.
47. Weick, K., *The Social Psychology of Organizing*, Addison-Wesley, Reading, Mass., 1979.
48. Weiss, C., Research for policy's sake: the enlightenment function of social research, *Policy Anal.*, 1977, 3, 531, 1977.
49. Winkel, G. and Sanoff, R., An approach to an objective analysis of behavior in architectural space, in *Environmental Psychology*, Proshansky, H., Ittelson, W., and Rivlin, L., Eds., Holt, Reinhart, & Winston, N.Y., 1970.

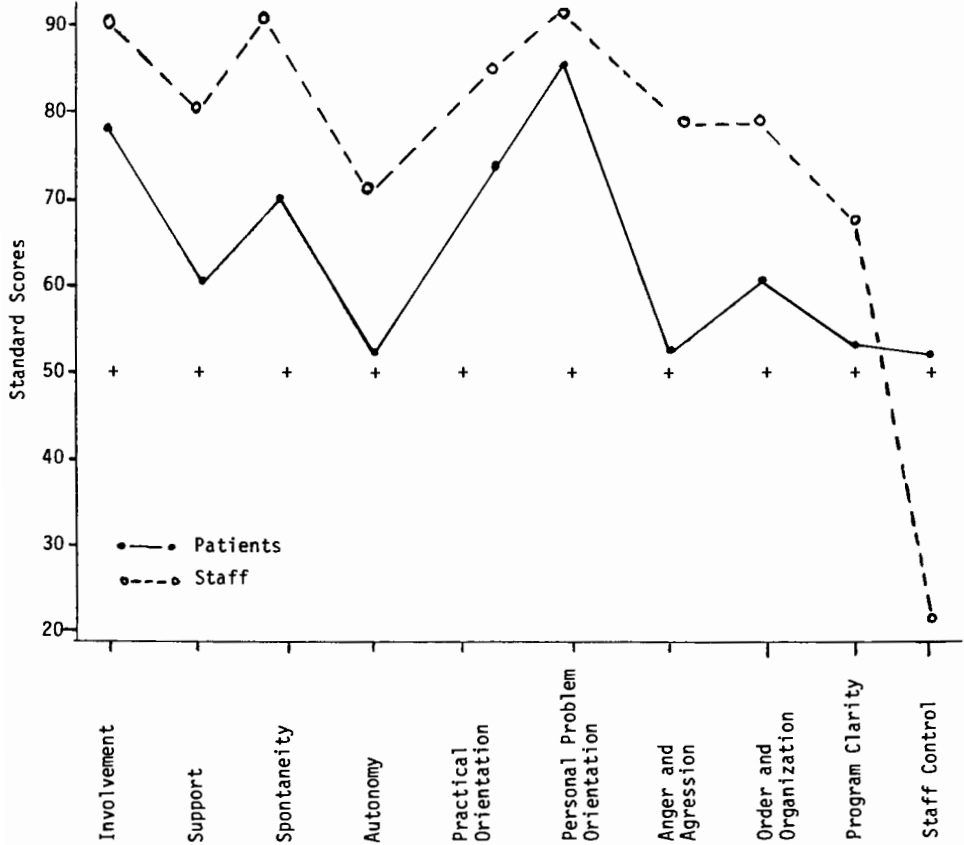


FIGURE 1. Example of profile generated by ward atmosphere scale.