A PORTABLE METHODOLOGY FOR ASSESSING PSYCHOSOCIAL ENVIRONMENTS IN A VARIETY OF SETTINGS

Austin T. DesLauriers

ABSTRACT

For more than a century, the social environment of a psychiatric treatment program has been considered a crucial factor in determining treatment outcome. During the last decade a number of instruments have been developed by Moos and his colleagues to systematically assess these social environment characteristics in a variety of settings. These instruments possess desirable psychometric properties and yield a series of subscale scores which can be categorized as reflecting three dimensions of the environment: 1) relationship dimensions; 2) personal development dimensions; and 3) system maintenance and change dimensions. The data produced from these instruments, which require a relatively short time to administer, can then be subjected to a variety of analyses - e.g. Hotelling's T^2, discriminant function analysis and profile analysis - to identify what areas of concern are shared by staff and members or serve to discriminate these groups. Such a methodology has many possible uses. A few of the possible uses include "Process" evaluation, monitoring program changes, and optimal matching of individuals and programs. The portability of the methodology derives from the facts that: 1) it requires only some access to "canned" computer programs and a minimal outlay for man-hours and materials; and 2) it has applicability to such a diverse range of settings as classrooms, correctional institutions, hospitals and community-oriented treatment programs.

It is ironic that in the two decades prior to the seventies, when behaviorism and its environmentally - oriented technology were in the ascendancy, so little effort was expended in trying to assess those characteristics of situations which would importantly influence behavior. To add to the irony, it was actually a shift away from a strictly response - reinforcement framework - toward an interactional position - which made the area of Environmental Assessment a salient one and has fostered its rapid growth in the '70's. Researchers, Clinicians and Workers throughout the field began to re-emphasize the importance of the person in the whole process of understanding behavior. Things have now progressed to the point that Bem and Funder (1978) in their very recent article in the Psychological Review on "Assessing the Personality of Situations" begin their article with the statement, "Behavior, as everybody knows, is a function of both the person and the situation." (Bem and Funder, 1978, p. 485). The key point here is that an important aspect of environmental assessment became: How does an environment, such as a therapy program, or a classroom, or a family prison, appear to most of the people in it: What dimensions, for example, are most salient? The consensus (Moos, 1975; Bem and Funder, 1978) is that a person's perceptions or impressions of a setting witll be an important determiner of his/her affective and behavioral response to that setting.

The issue, then, is to find some useful measures of the Psychosocial, or Perceived, Environment. Although many have been doing work in other areas of environmental assessment (see reviews by Frederiksen, 1972; Moos, 1973; and Pervin, 1978), two names have been mainly associated with psychosocial assessment: George Stern and Rudolf Moos and his associates. Stern's (1970) work has been primarily in educational settings, employs a complex methodology and statistics, and is not easily transportable to other kinds of settings. Consequently, with our interest in the use of this approach for program monitoring, planning and evaluation, we have become interested in the applications of Moos' (1975) Social Climate Scales which have been developed for a wide range of settings, are easy to administer and are fairly straight forward in their scoring, analysis and interpretation. In other words, they are very portable. In support of this, the Moos scales have been applied in environments ranging from total institutions (prisons, military companies) through semi-structured situations (hospital wards, community-based treatment programs, classrooms, university residences) to more or less naturally occurring and relatively less-structured settings (social groups, work milieus, families). Besides possessing acceptable psychometric properties (Moos and Otto, 1972), particular advantages of these measures for program or agency monitoring and evaluation are: 1) Low man-power costs for administration, scoring, etc; 2) By providing

an objective format in which all participants can furnish input about how the program is perceived, many of the inefficiences of group discussion - personality conflicts, power plays, inaccurate communications, time consumption, etc. - are avoided; and 3) Different program/agency issues can be addressed by using different instructional sets with the same set of questions, e.g. "Respond to these items in terms of how you see the program, agency, etc. now", or "Respond to these items in terms of how you would like the program to be if it could be just the way you'd like it."

So what do you have after you've administered and scored one of the social climate scales? What you have is 90 or 100 true-false answers which have been keyed to yield scores on 9 or 10 scales depending on the test used. You may have 2 sets of such scores for each person participating if you used both real and ideal instructional sets. The real score means can be converted to standard scores so as to compare your program with the norm group. Raw scores, from both real and ideal forms, can be used to compare groups within the agency/program. Real scores can be subtracted from ideal scores to yield a discrepancy measure of both satisfaction (or dissatisfaction) with the existing program and an indication of the desired direction of change. To aid interpretation of the group or individual profiles, the scale scores from any of their measures may be thought of as falling into 3 broad

categories: 1) Relationship dimensions, which identify the nature and intensity of personal relationships in the environment; 2) Personal development or treatment program dimensions which assess the basic dimensions along with personal growth and self-enhancement tend to occur in the environment; and 3) System maintenance and system change dimensions which assess the extent to which the environment is orderly, clear in its expectations, maintains control and is responsive to change. The relationship and system maintenance dimensions have been found to be relatively similar across environments while the personal development or treatment program dimensions have been found, as one would expect, to be particular to each environment assessed.

In order to understand how we have tried to extend the use of these measures, it is necessary to look briefly at what usual use has been made of the data from these tests. Typically, the scores have been divided into those belonging to staff and those belonging to members and then averages have been calculated for each group on each scale of the test. These averages are then translated into standard program scales, using the standard program scores for members only as the reference - the rational offered for this is that it will facilitate comparison of staff and member scores. Generally, the group means for the ideal forms are compared and, sometimes, the ideal minus real discrepancy scores are compared.

The problems we see with the methodology employed to date are three: 1) The eyeball technique used to compare profile graphs does not really tell us which group differences are fairly reliable and which are likely due to chance; 2) To compare two separate groups on the same norm group is not a very satisfactory way to understand what issues separate the two groups; and 3) We feel that these measures are very rich in information yield but that much of the information they contain is not being used.

What we are suggesting then, at a general level, is fairly simple-minded but simply has not been done yet: 1) The use of appropriate statistical techniques to help us determine both what group differences are reliable and where the cut-off point for significance occurs; 2) Using appropriate statistics to compare the groups directly rather than using an inappropriate norm group - it is hoped that the interpretation from this approach will also be more specific to the program: 3) Making greater use of the discrepancy scores; and 4) Down the road a bit, so to speak, making greater use of the expectation instructional set.

Let's take these suggestions to a more specific level. To do this, let's consider what you have and what you'd like to do with it. (What you have, as has been mentioned, are 9 or 10 scale scores for each person in your program/agency

who participated: Possibly you have 2 sets of these scores for each person. You also have the knowledge that these scale scores are moderately intercorrelated, i.e. they're not perfectly independent of each other and that quite likely the scores are grouped in some way, e.g. staff and members. What you will want to do with this information is determine if various kinds of differences are important or not, e.g. between your program/agency and the "Average", between various programs within your agency, between staff and members.

Given these conditions, the appropriate statistical approach is a multivariate one. Let me hastily add here - before everyone tunes me out completely - that what we are referring to is the use of standard "Canned" computer programs (all of our analyses were performed using the SPSS package) and simply being able to select those canned programs which will be of the most help to you in answering the questions of interest to you.

Let us move to a brief consideration of which statistical tools are most appropriate for which questions. As a first step you will wnat to perform either a MANOVA or a Hotelling's T^2. These are simply the multivariate analogues of the standard F-Test and T-Test: Which you use will depend on whatever you have your data assigned to 2 groups (Hotelling's T^2) or more than 2 groups (MANOVA). The question you are asking here is: Are there any overall differences between the groups

or, put another way, is it worth going any farther than this? If there is no significant difference between the groups at this level, then the advisability of further analysis is questionable since any differences subsequently found could be due to chance. Examples of the type of situation for which this analysis would be applicable are any of the group difference examples already cited or to comparisons across time, e.g. Is the program perceived any differently this year tha, say, last year? This kind of analysis can be used for real, ideal or discrepancy scores. If an overall difference is found, then it is appropriate to follow up with univariate analyses, F-Tests or T-Tests, to determine exactly where the significant group differences exist. Usually, the program automatically performs these analyses for you.

In conjunction with these analysis we recommend also running a discrimant function analysis. The basic question here is: What <u>set</u> of scales most contributes to group differences? In our scheme of things, we see this as serving two functions: 1) As a kind of a check on our univariate T-Tests or F-Tests - since those tests were not totally independent of each other there is the possibility of some capitalization on chance; and 2) More importantly, as an aid to interpretation - the scales which significantly discriminate the groups constitute a factor and it is worth trying to identify what overall concept ties those scales

together because it is that concept or issue which stands between the groups.

A final analytical tool, profile analysis, is mentioned in passing despite its relative analytical and interpretive complexity because it has been constructed specifically to aid in the discovery of similarities and differences in profile data. The basic analysis involves utilizing the MANOVA program in specific ways so as to discover what patterns with regard to profile shape, elevation or scatter distinguish the groups (applicable to only 2 groups at a time). Although we will not be presenting examples of profile analysis today, we have used it on the W.O.T.C.H. data and our basic rule has been to include it in our findings when its meaningfulness is fairly clear and it will serve as an aid to our overall interpretation.

Let us assume that you have collected data on the psychosocial environment of your program/agency and subjected it to the appropriate analysis. How can your program/agency best make use of the available information in its program evaluation, planning and monitoring? Let us further assume that the data you have collected and analyzed represents your first use of these instruments so you don't really have anything to compare your results with and are faced with the somewhat intimidating question: What does it all mean? You will probably have analyzed 3 sets of scores for your groups,

so let's look at each set very quickly. The Real scores are quite straightforward: They are a measure of how the program is currently perceived and group differences simply represent differences in present perception of the program. The questions to be dealt with, then, will be: 1) Are the group differences logical and expected ones in view of the goals and design of the program; and 2) Are the group differences in support of or in opposition to the goals of the problem. For the ideal scores, the groups are indicating how they would wish their program to be and the question for evaluation then becomes how do these ideals, these goals to be striven for, compare with the program/agency's stated goals? Finally, the discrepancy scores (Ideal minus Real) present 2 pieces of information for the decison-making process: 1) The level of satisfaction with the existing program; and 2) The direction of desired change. It would seem logical that over the meaning of the information for the program/agency had been determined, decisions could have made as to how one would like the present perceptions, ideals, level of satisfaction or direction of desired change to be different (Evaluation) and what would be the program changes most likely to accomplish this (Program Planning). Changes could then be implemented and monitored - by exactly the same methodology already described - as desired. The cycle as we see it, then, would involve 5 steps: 1) Collecting

and analyzing the data; 2) Interpreting its meaning for the program; 3) Evaluating those changes, if any, which would be desirable; 4) Planning and implementing the changes; and 5) Monitoring the program changes, i.e. collecting and analyzing new data.

Since these instruments - the social climate scales - are a largely untapped source of program information, I think it would be worth mentioning some additional possible uses of this methodology, most of which have never been attempted as far as I can determine. Moos has suggested some of these, others have occurred to us in process. One could involve people outside the program/agency as a comparison group - Moos suggests asking observers in to rate the program, but they could also rate written descriptions of the program. There is the whole question of matching clients with the treatment program most appropriate for them. Here the use of the expectation form of the instruments might be quite useful: Do those clients do best in the program whose expectations are most congruent with actual psychosocial environment or the program (Otto and Moos, 1974)? Two other approaches to this issue are also possible: 1) Whenever data is collected, clients may be grouped according to how well it is believed they have made use of the program and their average profiles compared; or 2) If a number of agencies in an area are collecting this kind of data, a

pooling of environmental information could be accomplished with the aim of making optimal referrals - this would seem to be an important step toward the elimination of wasted time and lost persons in the mental health system. A related possible use of this kind of information is to simply improve the accuracy of program descriptions based on the perceptions of those actually involved in the program. A final possible use of this kind of data which I shall mention here is the designing of research with practical relevance for busy program staff.

In summary, we can say that quite often program descriptions, monitoring, planning and evaluation are carried out largely by the agency/program head(s) with some input from those considered expert sources or perhaps including some informal observation of day to day functioning "Just to see how things are going." What we have suggested, we feel, is an economical way in which to include two other valuable sources of information into the decision-making process. The first source of information is from the Moos scales. This is really information from the program/agency participants but I emphasize the Moos scales to underscore the facts that in this form: 1) Many of the inefficiencies of group discussion - mentioned earlier - are avoided while much of the richness of information which can come from such discussion is preserved; and 2) The information obtained is

objective in the sense that all participants are giving their impressions of the program/agency along certain fairly specific dimensions. The second source of information we would add to the decision-making process comes from statistics and we see this as helping to answer two questions: 1) Are the group differences reliable; and 2) Where do the reliable differences occur? At this point, I would like to note that, ultimately, I think this is the least important source of information in the process so, if one doesn't have access to a computer, there would still seem to be considerable value in adding the information from the social climate scales to the evaluation process.

Note that the locus of responsibility for program/agency decisions does not change here - the same program head, or staff group, or staff-member committee - still must make the important evaluative judgements and planning decisions. They still must utilize what is the critical source of information in the process - their own knowledge of the program goals, of the politics of the situation and the consequences involved, of the costs likely to be incurred, their perceptions of the capabilities of the people they work with, etc. So there is still plenty of room for what Gerry Stone at Western calls "Blue-Sky" discussion.

I will conclude on a note of levity with a point. I think we all get a little nervous when the realm of science, often associated with the Ivory Tower of Academia, and especially that symbol of science, the computer, threatens to alter some of our preconceived notions, our sense of how it is "in the field". It's worse if the threat is perceived as a takeover of our decision-making powers in some area. Well, we can all calm down because in this case science, Dr. Moos and the computer are simply providing us with what could be very useful information. It will in no way compel us to act against our will and, in fact, for some of you, it may provide very useful ammunition against that smart alec at the staff meetings who keeps pushing for sweeping changes. You can hit him with: "I hear what you're saying, Fred, but if you'll just have a look at this print-out from our DATA BANKS blah...blah...blah." How can he rebut the print-out, especially when he doesn't really know what it says?

So, rest easy, folks, our attempt has simply been to help us all print the pictures a little more clearly but the process of program monitoring and evaluation remains, with the rest of psychology, an art form.

References

Bem, D.J. and Funder, D.C. "Predicting More of the People More of the Time: Assessing the Personality of Situations," Psychological Review, 1978, 85, 485-501.

Frederiksen, N. "Toward a Taxonomy of Situations." American Psychologist, 1972, 27, 114-123.

Moos, R.H. The Social Climate Scales: An Overview. Palo Alto: Consulting Psychologists Press, Inc., 1975.

Moos, R.H. "Conceptualizations of Human Environments." American Psychologist, 1973, 28, 652-665.

Moos, R. and Otto, J. "The Community-Oriented Programs Environment Scale: A Methodology for the Facilitation and Evaluation of Social Change." Community Mental Health Journal, 1972, 8, 28-37.

Otto, J. and Moos, R. "Patient Expectations and Attendance in Community Treatment Programs." Community Mental Health Journal, 1974, 10, 9-15.

Pervin, L.A. "Definitions, Measurements and Classifications of Stimuli, Situations and Environments." Human Ecology, 1978, 6, 71-105.

Stern, G. People in Context: Measuring Person Environment Congruence in Education and Industry. New York:Wiley, 1970.

Dear Mr. Toch:

This is a more or less orderly collection of the notes I used for my paper. I hope it is of interest.

Cordially,

Austin T. DeLamater