DOCUMENT RESUME

ED 252 761 CG 017 929

AUTHOR Friedlander, Myrna L.; And Others
TITLE A Behavioral Analytic Model for Evaluating Counselor Training Programs.
PUB DATE Aug 84
NOTE 10p.; Paper presented at the Annual Convention of the erican Psychological Association (92nd, Toronto, tario, Canada, August 24-28, 1984).
PUB TYPE Reports - General (140) -- Speeches/Conference Papers (150)

EDRS PRICE MF01/PC01 Plus Postage.
DESCRIPTORS Behavioral Science Research; *Counselor Training; *Measurement Techniques; Models; Practicums; *Program Effectiveness; Program Evaluation; Psychologists; *Student Evaluation

ABSTRACT

This paper introduces a rationale and method for constructing behavioral analytic measures of training program effectiveness that can be adopted by directors of training in diverse settings. The model includes derivation of problematic, on-the-job situations and effective responses from a target sample of trainees and builds an evaluation measure based on this derivation. The target population consists of entry level practicum students in their second year of doctoral training. The model contains five steps. In the first step, "situational analysis," a survey of the relevant characteristics of the environment is conducted with a sample of subjects currently performing in that setting. These subjects generate a detailed list of problematic situations that they have encountered while performing on the job. The next step, "response enumeration," is a sampling of the target population's common responses to these situations. The following phase, "response evaluation," uses a panel of experts to evaluate the effectiveness of the various responses to the problematic situations generated in the preceding phase. These first three steps in the process are the "criterion analysis." The next step is to construct a format (such as the Practicum Evaluation Measure) for presenting the selected situations (plus possible responses) to successive samples. The final step is to evaluate the measure using standard psychometric procedures. Details of ongoing research with this model at a psychology department training site are presented. Several program evaluation designs are recommended, to assess the effectiveness of a program for trainees with different characteristics, different supervisory formats, or different training components. (LLL)

A Behavioral Analytic Model for Evaluating Counselor Training Programs

Myrna L. Friedlander, Linda G. Ward, and Howard Ferrin

State University of New York at Albany

Paper presented at the annual convention, American Psychological Association, August, 1984, in Tororto, Canada.

A Behavioral Analytic Model for Evaluating

Counselor Training Programs

Abstract

Although on-site training is presumed to be an effective preparation for professional psychologists, existing measures reflect global characteristics of counselors rather than the degree to which trainees have mastered specific competencies. This paper introduces a rationale and method for constructing behavioral analytic measures of training program effectiveness that can be adopted by directors of training in diverse settings. Details of ongoing research with this model at a psychology department training site are presented. Several program evaluation designs are recommended, to assess the effectiveness of a program for trainees with different characteristics, different supervisory formats, or different training components.

Practicum and internship training are integral components of the professional preparation of counseling psychologists. Although on-site training is presumed to be an effective preparatory method, existing measures of effectiveness relect global characteristics of counselors (e.g., facilitative conditions) rather than the degree to which trainees have mastered specific competencies needed for their professional roles. The purpose of this paper is to introduce a model for constructing behavioral analytic measures of training effectiveness, a model that can be adopted by directors of training and supervisors in diverse settings to evaluate their program's effectiveness. The paper (a) presents the rationale and procedure for developing such measures, (b) describes our ongoing research with the model at a psychology department training site, and (c) recommends several different program evaluation designs with behavioral analytic measures.

The need for a behavioral analytic approach was suggested by our observation of the increased diversity of training agencies and the lack of theoretically based evaluation criteria. Lambert (1980) proposed that researchers endeavor "to identify in a prescriptive sense the ideal learning environments for given students at particular times" (p. 443). This prescription implies a score of potential criteria of effectiveness. While current theoretical models of training and supervision (cf. Loganbill, Hardy, & Delworth, 1982; Stoltenberg, 1981) provide general guidelines for matching supervisory approach to trainees' needs, there are few conceptual guides for constructing a training program to enhance counselor development. In addition, a delineation of suitable outcome criteria for training and supervision has been notably lacking in the literature. Empirical studies have, of necessity, taken a global approach to the problem, e.g., assessing the effects of supervisory influence on "professional

and personal development" (Friedlander & Snyder, 1983; Heppner & Handle , 1981. There is, however, a considerable gap between these broad criteria and the specific tasks that trainees must master in order to perform effectively in a particular setting. A college counseling center may expect trainees to develop minority group programs, for example. Suitable evaluation of training and supervision at this center requires an assessment of trainees' ability to develop and evaluate programs as they confront various predictable problems. These skills would be irrelevant in another setting -- a child guidance center, for instance, where trainees must learn how to consult with school psychologists, teachers, and parents.

Even if existing theoretical models of counselor training suggested specific criteria for trainees at different levels, there are inherent proolems in designing outcome measures based solely on theory. To do so entails the following assumptions (Goldfried & Kent, 1972): (a) the set of principles represented by the theory provides a comprehensive picture of the target population, (b) participants' responses are not subject to environmental variability, therefore (c) the evaluator need not attend to situational or population-specific sources of variation. Clearly these assumptions would be violated in attempting to develop a sound instrument for evaluating training programs across settings. Due to the nature of the profession, we cannot afford to overlook the potentially confounding effects of trainees' attributes as they interact with the training program and client population. For example, some personal characteristics may be unsuited for a setting where the trainee has little autonomy but major responsibility for counseling severely disturbed clients.

Given the theory-practice lag and the diversity of trainees, clients, and settings, one might be tempted to abandon the search for relevant criteria of program effectiveness. If we consider a training program as an intervention, the ideal set of criteria for assessing the effectiveness of this intervention would be (a) theoretically derived, (b) relevant to the existing program and (c) trainee population yet (d) sensitive to individual differences. While not all of these requirements can be met simultaneously, the behavioral analytic model (Goldfried & D'Zurilla, 1969) is a promising vehicle for designing population-specific measures of training program effectiveness.

The Model

The salient question becomes, "What constitutes a relevant outcome for this trainee population?" Our approach to assessing a counselor training "intervention" is an adaptation of Goldfried and D'Zurilla's (1969) behavioral analytic model for evaluating competence. "Competence" is defined operationally as "the effectiveness or adequacy with which an individual is capable of responding to the

various problematic situations which confront him" (Goldfried & D'Zurilla, 1969, p. 161). The behavioral analytic approach to assessment emphasizes both individuals and situations as well as specific behavior-environment interactions. This procedure reflects an attempt to maximize individual and situational differences and to minimize the potential bias of pre-existing theory.

The model includes derivation of problematic, on-the-job situations and effective responses from a target sample and builds an evaluation measure based on this derivation. The model contains five steps. In the first step, "situational analysis," a survey of the relevant characteristics of the environment is conducted with a sample of subjects currently performing in that setting. (One assumes that this first sample adequately represents the target population.) These subjects generate a detailed list of problematic situations that they have encountered personally while performing on the job. The next step, "response enumeration," is a sampling of the target population's common responses to these situations. The following phase, "response evaluation," uses a panel of experts to evaluate the effectiveness of the various responses to the problematic situations generated in the preceding phase. These first three steps in the process are the "criterion analysis." The next step is to construct a format for presenting the selected situations (plus possible responses) to successive samples. The final step is to evaluate the measure using standard psychometric procedures.

An Illustration: Psychological Services Center

The following example illustrates how the model may be adapted to assess the effectiveness of a doctoral practi training experience.

Step 1: Define the Intervention

First, the training intervention needs to be described in terms of goals, objectives, and procedures. In our example, the intervention consists of one year of supervised practicum at a psychology department training site (Psychological Services Center; PSC) at a northeastern state university. Clients from the urban community come on a fee-for-service basis (sliding scale), and the clientele represent highly diverse life circumstances and presenting problems. The PSC is staffed by a director (a licensed psychologist) and a full-time secretary. A number of faculty members provide supervision, both individual and group, and opportunities are available for live observation and audio- or videorecording. Second year doctoral students typically carry a caseload of from 4 to 6 clients (individuals, couples, or families) from September through May. In addition, trainees are also responsible for handling telephone intakes, walk-in crisis intervention cases, and (occasionally) formal psychological evaluations from local agencies.

The general goal is to prepare doctoral students in counseling and clinical psychology for a full-time internship at an APA-accredited site. Specific objectives, in the form of "minimal competencies," were developed by an appointed committee of counseling psychology faculty two years prior to the beginning of this research. These objectives fell into several categories: assessment, case management, interviewing skills, treatment planning, and follow through.

The training intervention has as its primary objective to develop and enhance basic counseling skills in these areas and, in so doing, to influence trainees' expectations of self-efficacy (Bandura, 1977, 1982) related to these skills. The assumption is that the first stage in counselor training is developing a sense of one's competence in the professional role of counselor (cf. Stoltenberg, 1981). While evaluation of the other objectives in this training program (such as actual skill attainment) could be assessed within the behavioral analytic model, for the purpose of this illustration, only the objective of enhancing trainees' self-efficacy expectations is considered.

Step 2: Define the Population

The target population consists of entry level practicum students in their second year of doctoral training in an APA-accredited counseling psychology program at the State University of New York at Albany. Prior to practicum, some students have completed only a semester-long prepracticum experience, while other students (having entered the program with a master's degree) have had some previous supervised counseling experience.

Step 3: Conduct a Criterion Analysis

The initial step was to generate a series of problematic situations confronting the target population. In our example, the domain of situations was limited to actual counseling and assessment skills (i.e., excluding peer interactions of counselor-supervisor relationships).

A thought listing procedure was adopted. Subjects ($\underline{N} = 6$), the group of students completing practicum in May, 1983, were solicited individually during the last week of the Spring semester. Participants were asked to generate the problematic situations that they had encountered personally during the past year. Seven general categories were constructed in order to provide subjects with some guidelines for organizing their thoughts: client assessment and conceptualization, interviewing skills (managing the flow of the session), planning and carrying out treatment, technical skills, managing the client/counselor relationship, case management, miscellaneous.

Subjects described a total of 112 situations, in each of these seven categories. For each situation they also indicated (a) whether they had received supervision about the problem, and (b) confidence in their ability to handle the situation should it arise again, on a 0 (least) to 9 (most) scale. These confidence ratings were our adaptation of Goldfried and D'Zurilla's (1969) phases of response enumeration and evaluation (cf. Phillips, 1983). The "expert" judgment about the "effective" response to each situation was, in effect, the trainees' own assessment of his or her self-efficacy. These ratings allowed us to identify the range of "effective" responses in the initial sample.

Step 4: Develop a Measurement Format

Having generated the domain of problematic situations and trainees' responses to them (i.e., their self-efficacy ratings), we proceeded to develop a uniform measurement for subsequent administrations in the target population. First, we reviewed the thought-listed situations and constructed items to reflect the most common problems. These were then entered into a new format. The resulting instrument, the Practicum Evaluation Measure (PEM) contained 20 items. Subjects rate their "confidence in (their) ability to..." on a 0-9 scale (not confident to completely confident).

Step 5: Evaluate Psychometric Characteristics

Because the pool of items was drawn from a small sample, we conducted an item analysis on a second sample (PSC practicum trainees in the 1983-84 academic year) from the target population. This was to insure that the problematic situations generated in Step 3 were applicable to successive groups of trainees. To do this, half of the entering practicum students (n = 5) in Fall, 1983 (chosen randomly) completed the PEM in early September. Means on each PEM item in this pre-measure were computed, and 3 items were eliminated whose M 7.0. (These eliminated items were considered to represent situations that were not particularly problematic for the second sample.)

Second, all practicum students (N = 10) completed the 17-item PEM in May, 1984, at the end of their training year. These data were used to provide an estimate of the internal consistency of the measure. Interitem reliability was estimated at .83. Additionally, an indication of its sensitivity to pre-testing was determined by an F test of the difference between groups who (a) had completed the instrument both pre- and post- versus (b) those who had completed only the post-test. Results were nonsignificant, $F(1, 8) = 3.59$, ns.

Finally, additional evidence of the validity of the PEM will be determined in the following manner. Pre-/post- data will be collected annually until 1987, such that the final validation sample will be at least 40. These subjects' post-test scores will be compared with a

second measure, the Self-Efficacy Inventory (S-EI; Friedlander & Snyder, 1983). In contrast to the PEM, the S-EI is a global index of counselor self-efficacy. The S-EI contains 21 items reflecting completion of academic requirements, assessment, individual, group, and family counseling, and case management. Like the PEM, trainees indicate their confidence in their ability to perform these activities on a 0-9 scale (not confident to completely confident). The S-EI has an internal consistency reliability of .93, and in previous research (Friedlander & Snyder, 1983) it was significantly correlated with level of training. Items from the PEM and S-EI are randomly combined in order to minimize a potential response bias. A significant positive correlation between the two measures will indicate concurrent validity, since self-efficacy expectations of global skills (S-EI) and of situation-specific competencies (PEM) should be related.

Recommendations

The nature of additional tests of reliability and validity of behavioral analytic measures depends on the researcher's aims. Parallel forms of the instrument might be devised from the situations generated during the criterion analysis, for example. Parallel form reliability estimates could be obtained, and the use of two forms would decrease sensitivity to pre-testing. Concurrent validity could be established by correlating the behavioral analytic measure with supervisors' ratings of their trainees' competencies. An additional test of validity would be a comparison of the post-test responses of comparable trainees from different settings. If the measure is valid, trainees in setting A (the one for which it was designed) should score significantly higher than trainees in setting B.

Campbell and Stanley (1963) have provided examples of experimental and quasi-experimental research designs that might be tailored to researchers' individual needs. They also discuss threats to internal validity (history, maturation, testing, instrumentation, statistical regression, selection biases, experimental mortality, and selection-maturation interaction) and external validity (reactive effect of testing and/or experimental arrangements, the interaction of selection bias and the experimental variale, and multiple treatment interference) that the naturalistic researcher should be familiar with when choosing a particular design.

With these threats to validity in mind, we suggest several program evaluation designs for use with behavioral analytic measures. As one example, in order to assess the effectiveness of a program for trainees with different characteristics (age, sex, previous counseling experience), a factorial design could be used. The viability of this design is, of course, limited by the small numbers of students typically involved in training programs.

Second, it might be of interest to contrast the training effectiveness of different supervisory formats (e.g., individual versus group versus co-counseling, or live versus no observation versus audio- or videorecording). A time series design with multiple Ns of 1 could be employed. With this design, each trainee is tested at specific intervals and the supervisory format of interest is introduced at a point chosen randomly. Although full experimental control is lacking in this quasi-experimental design, it can be used effectively for program evaluations despite its limitations. The time series provides strong control over sources of internal invalidity. It is possible, however, that change-producing events other than the supervisory technique of interest may occur, diminishing the researcher's confidence in the effectiveness of the supervisory intervention. Results will be specific to each trainee and not generalizable to all. However, an N of 1 approach may be useful in tracking the growth of an individual trainee over time, and multiple Ns of 1 would be more reliable.

Another research possibility with a heterogeneous population is a pre-experimental design using one group in a pretest--training intervention--posttest situation. This design minimizes internal invalidity with regard to selection bias but fails to provide control over such factors as the effects of pre-testing and the effects of factors other than the training intervention that might occur between measurements. Without a control group, it would be difficult to rule out alternative plausible hypotheses for the measured effects of the training intervention. Ths design may be useful, however, when only one group of trainees is available.

When comparison with a control group is feasible, several other experimental designs may be warranted. Different elective components of a program could be assessed by using the trainees who do not participate in a given rotation as controls. A post-test only control group design can be used in situations where randomization is not possible, when pre-tests are inconvenient or highly reactive, or when the trainee's anonymity is an issue. This design provides strong controls for sources of both internal and external validity. Finally, with a pre-test/post-test control group design, the pre-test is used as a covariate. This design also provides stong controls over sources of internal and external validity. With the pre-test as a covariate, invalidity due to the the interaction of selection bias and the training intervention can be minimized.

Although each of these designs has limitations, our intent is to suggest possible training evaluation designs to use with behavioral analytic measures. The cumulative results of such program evaluations eventually may provide directions for refining theoretical models of counselor training and supervision.

References

Bandura, A. (1977). Self-efficacy: Toward a unifying theory of behavioral change. Psychological Review, 84, 191-215.

Bandura, A. (1982). Self-efficacy mechanism in human agency. American Psychologist, 37, 122-147.

Campbell, D. T., & Stanley, J. C. (1963). Experimental and quasi-experimental designs for research. Boston: Houghton Mifflin.

Friedlander, M. L., & Snyder, J. (1983). Trainees' expectations for the supervisory process: Testing a developmental model. Counselor Education and Supervision, 22, 342-348.

Goldfried, M. R., & Kent, R. N. (1972). Traditional versus behavioral personality assessment: A comparison of methodological and theoretical assumptions. Psychological Bulletin, 77, 409-420.

Goldfried, M. R., & D'Zurilla, T. J. (1969). A behavioral analytic model for assessing competence. In C. D. Spielberger (Ed.), Current topics in clinical and community psychology (Vol. 1). (pp. 151-193). New York: Academic Press.

Heppner, P. P., & Handley, P. G. (1981). A study of the interpersonal influence process in supervision. Journal of Counseling Psychology, 28, 437-444.

Lambert, M. J. (1980). Research and the supervisory process. In A. K. Hess (Ed.), Psychotherapy supervision: Theory, research, and practice. (pp. 423-450). New York: Wiley.

Loganbill, C., Hardy, E., & Delworth, U. (1982). Supervision: A conceptual model. The Counseling Psychologist, 10(1), 3-42.

Phillips, S. D. (1983, March). Developing evaluation measures for career interventions. Paper presented at the American Personnel and Guidance Association annual convention, Washington, D. C.

Stoltenberg, C. (1981). Approaching supervision from a developmental perspective: The counselor complexity model. Journal of Counseling Psychology, 28, 59-65.