

An evaluation research design was developed as an attempt to provide a more satisfactory approach to microcounseling training program evaluation. Trainee performance was measured three times during a counseling practicum, with microcounseling training occurring between the second and third observations. Trainee performance was compared to a predetermined standard for counselor behavior. Results were analyzed for both the differences between observations, and the degree of similarity to the model. Counseling behavior of trainees after microcounseling training was significantly different from their behavior prior to the training. After training they were more like the standard. The trainees performed less like the standard after some counseling experience, but before receiving microcounseling training.

EVALUATING MICROCOUNSELING TRAINING

Comparing Trainee Performance to a Predetermined Standard for Counseling Behavior

STEPHEN JOEL GILL

CARL F. BERGER

University of Michigan

GEORGE L. COGAR

Veterans Administration

Tomah, Wisconsin

Kasdorf and Gustafson (1978), in the discussion of the state-of-the-art of microcounseling research, suggest that microcounseling training is both viable and in need of improved evaluation models. The microcounseling approach to counselor training teaches interviewing skills through a process of observing one's performance on videotape. Specific behaviors, helpful to counselor-client interaction, are systematically taught based on learning principles. These skills have been identified by investigating the interdisciplinary research on effective interviewing and by a trial-and-error approach to videotaping examples of the skills (Ivey and Authier, 1978).

EVALUATION REVIEW, Vol. 7 No. 2, April 1983 247-256

© 1983 Sage Publications, Inc.

0193-841X/83/020247-11\$1.35

Researchers have reported success in teaching single skills to guidance support personnel (Haase and Dimattia, 1970), to graduate students in a counselor education program (Ivey et al., 1968), to graduate students in a clinical psychology program (Moreland et al., 1970), to high school advisors (Kerrebrock, 1971), to residence hall counselors (Scroggins and Ivey, 1976), and to psychiatric nurses (Authier and Gustafson, 1976), among others.

The inherent problems of any program evaluation plague these studies, raising questions regarding the validity of their findings and the ability of counselor training programs to assess accurately the impact of their methods. The microcounseling approach facilitates objective evaluation through its emphasis on observable, measurable counselor behaviors. Investigators, however, have tended to rely on commonly used evaluation designs that do not take full advantage of the specificity of outcomes, nor do these studies adequately control for the effects of variables other than the training method.

Microcounseling training evaluation studies typically use either a one-group, pretest-posttest design or a pretest-posttest control group design. Campbell and Stanley (1963) have described each of these approaches and have cautioned investigators regarding the weaknesses of each design, particularly the one-group, pretest-posttest design. This method cannot usually reject a number of rival hypotheses that may be more responsible for change in the participants than the training itself. These alternative explanations hold that change could be due to history—events occurring during the course of the experiment but extraneous to the training procedures, such as nontraining counseling experiences and intragroup interactions of trainees; maturation—the normal course of development for any individual; testing—practice that occurs from performing the skill for each experimental observation; or, statistical regression—the tendency of extreme scores to change as a result of the statistical properties of measuring the same group twice.

Pretest-posttest measures tend to be an examination of gross differences that often fail to answer two vital research questions: "Was the change in the direction expected by the researchers?" and "Was the change large enough to be considered significant?" Post hoc analysis may be used to answer the first question, but not the second. Even a significant difference in the expected direction may not produce the desired change that the researchers predicted.

The pretest-posttest control group design, when assignment of subjects to the experimental and control groups is truly random, can reject

the rival hypotheses described above. In this method, counseling trainees are randomly assigned to the two groups, skills of all trainees are measured, the experimental group receives the training, and then both groups are measured again for comparative group change. Hersen and Barlow (1976) have described some of the difficulties occurring with this design: the ethical dilemma of withholding treatment from some subjects; the practical problems of having enough subjects for comparison groups; and obscuring important individual change by comparing group means only. Other problems with the pretest-posttest control group design include the contamination that can occur when trainees from both groups interact with each other outside of the training program; and the fact that trainers have not had the resources or knowledge base to evaluate their programs adequately (Nuttall and Ivey, 1978).

METHOD

To alleviate these evaluation problems, a measurement technique and research design rarely used by counselors were applied to the evaluation of a microcounseling training program. First, the two counselor educators involved with the study developed a description of the seven counseling skills they considered important to use during the first five minutes of a typical counseling session. The two investigators independently identified an "ideal" frequency for each skill, and then agreed on a model of performance which includes all of the skills (see Table 1). The trainees' performance could then be compared to this standard for counselor behavior. In this way, significant differences could be detected not only by absolute changes in observed counseling behavior, but also by the direction of those changes in relation to the standard for counseling behavior.

Next, a research design was developed as an alternative to the simple pre- and posttreatment design and the separate experimental and control group design. The interrupted time-series approach was adapted to a counselor training situation. Cook and Campbell (1979) have discussed the advantages of this design. The modified time-series in this study used the following three steps: (1) counseling with observation was followed by a period of time with no treatment; (2) counseling with observation occurred again and was followed by trainee participation in the training program; and (3) counseling with observation occurred a

TABLE 1
 Predetermined Standard for Performance of Microcounseling Skills
 During First Five Minutes of an Exploratory Counseling Session

<i>Microcounseling Skill</i> ¹	<i>Skill Frequency as Percentage of Total Counselor Responses</i>
1. Closed Question (CQ) inquiries which require only a "yes" or "no" answer or other short, non-exploratory response	3
2. Open Question (OQ) inquiries using words such as "how" or "what" which require clients to express themselves in an extended verbal response	25
3. Minimal Encouragers (ME) non-verbal and verbal leads which seek to elicit continuation of the client's response (head nods, um-hmmm, etc.)	20
4. Reflection of Feeling (RF) verbal communication which relates only the emotion of what the client's message conveys	13
5. Reflection of Content (RC) verbal communication which relates only the meaning of what the client says	12
6. Paraphrase (P) verbal communication which restates what the client has said using different words	15
7. Selective Focus (SF) verbal communication which is limited to a narrow portion of the clients message	2
8. Other (O)	<u>10</u>
Total	<u>100</u>

1. These skills and their definitions were adapted in part from the microcounseling training program of Ivey and Gluckstern (1974).

third time. By not dividing into separate groups, this design helped to alleviate several problems found in other designs: not having enough subjects to split into two groups; having to withhold treatment from one group; and risking that subjects in an experimental group would somehow affect the performance of a control group through their contact.

A repeated measures analysis of variance statistical technique (Harris, 1975; Hersen and Barlow, 1976) was used to detect differences among observation measures. Planned, correlated t-tests (Ferguson, 1976) were used to detect if these differences were due to experience or treatment effects. If experience is a major variable, differences between observation one and observation two should be detected. If a treatment effect exists, differences between observation two and observation three should be significant. The authors hypothesized that the counseling behavior of trainees following a microcounseling training program would be significantly different from their behavior at the two observations prior to the training, and that after training the trainees would be more like a theoretical model of counselor behavior than they were prior to the training.

PROCEDURE

Twelve students in one section of a practicum course within a Guidance and Counseling M.A. degree program participated in the microcounseling training. These trainees ranged in age from 23 to 36. There were 10 female and 2 male students. All of the trainees counseled the same coached client with the same presenting problem for five minutes at three different times during the course. Three weeks separated each set of counseling sessions, placing the observations at the beginning, middle, and end of the term. The client was a 30-year-old female doctoral student who was asked to role-play a problem situation. The trainees did not know the client nor were they aware of the specific behaviors being observed. All sessions were videotaped through an observation window. Due to absences, only 10 counselors-in-training completed all three videotaped counseling sessions.

Prior to the third observation, the trainees viewed and discussed videotape models of a counselor performing examples of the seven different skills (see Table 1). Each videotape was approximately 2 to 3 minutes in length and showed a female counselor helping a male client.

Precautions were taken to insure that in-class experiences were the same for all trainees and that there were no significant events that would be a threat to internal validity.

Videotape of the counseling sessions was observed by two experienced counselors trained in the identification of microcounseling skills. They observed the first five minutes of each session and achieved consensus on the frequency of the behaviors modeled during the training program.

RESULTS

Performance frequencies for each skill were converted to a percentage of total number of trainees' responses made during the session. This was done to reduce the likelihood that quantity of a particular trainee's responses would influence the study. An evaluation of the total trainee group's performance was determined by comparing these percentages to the standard presented in Table 1. Total absolute percentage differences between actual performance of all skills and the standard are presented in Table 2. This table also presents the means and standard deviations of absolute percentage differences between the actual and standard for the total trainee group at each observation. An examination of the total absolute percentage differences between the actual performance and the standard for performance at each observation time appears to indicate that after training the students were more like the standard than before training.

An analysis of variance (ANOVA) of the means for the total group at each observation time indicates that there were significant differences ($p = .031$) between one or more observations (see Figure 1). Also the ANOVA results indicated strong differences ($p = .004$) among counselors within the observation times.

The planned correlated *t*-test indicates that a strong difference ($p = .007$) exists between observation two and observation three. An examination of the means indicates that there was a definite decrease in the difference between trainees' actual scores and the standard frequencies from observation two to observation three. The difference between observation one and three is not significant ($p = 0.75$). Student performance after training (time three) is more like student performance before experience only (time one) than it is like student performance after experience but before training (time two).

Since these findings may be due to heterogeneity of variance (see Figure 1) rather than because of little change in each student's perfor-

TABLE 2

Total Percentage Difference Between Trainees' Actual Skill Performance and the Standard for Skill Performance at Each Observation Time

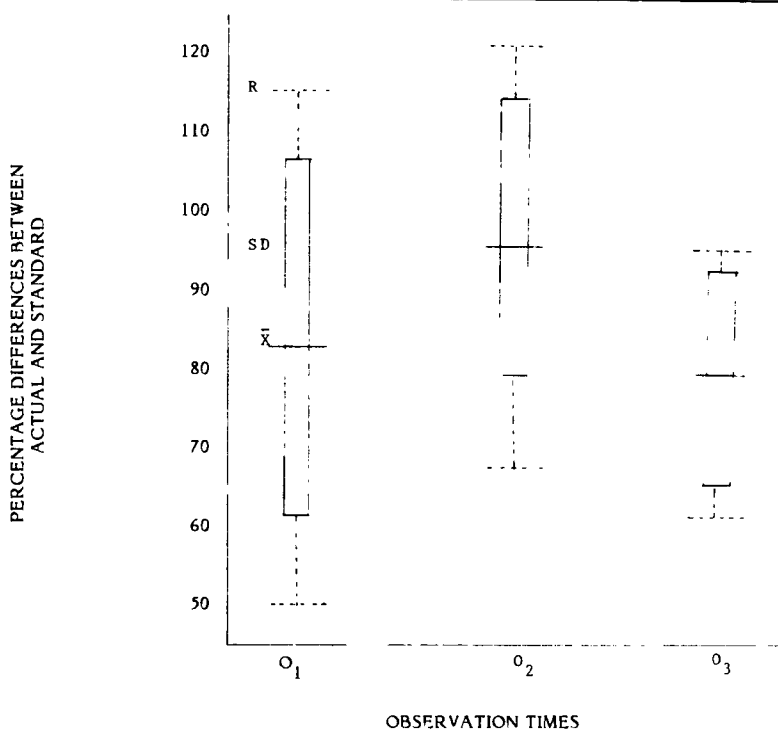
Trainee (<i>N</i> = 10)	<i>Percentage Differences Between Actual and Standard for Skill Performance at Each Observation Time</i>		
	θ_1	θ_2	θ_3
A	50.0	101.6	61.6
B	62.2	89.6	91.1
C	101.1	85.3	78.8
D	88.0	106.2	91.4
E	114.7	112.8	94.0
F	85.9	84.0	80.2
G	62.4	73.4	72.6
H	63.5	68.0	64.0
I	88.6	108.0	63.2
J	115.0	120.0	94.0
Total	831.4	948.9	790.9
\bar{X}	83.1	94.9	79.1
SD	23.0	17.4	13.2

mance, the Cochran test for equality of variance (Marascuilo, 1971) was used. This test yielded a *C* value of .526, which is not significant at the .05 level. Therefore, the variances are assumed to be equal, even though there is an apparent decrease in variance from observation one to observation two.

The *t*-test also shows that there is a significant difference at the .05 level between observation one and observation two ($p = .039$). Average differences between actual performance and the standard indicates that the trainees were less like the standard after experience only.

CONCLUSIONS

The hypotheses of the authors were partially supported by the results of this study. The counseling behavior of trainees following a micro-counseling training program was significantly different from their behavior preceding the training. This change was in the direction that the authors expected and is statistically significant. In fact, there is a dramatic change toward the predetermined standard for 9 of the 10 trainees. This would not be expected if the intervention had no effect.



Source	Sums of Squares	d.f.	Mean Squares	F	P
Group Observation	13.47	2	6.74	4.25	.031
Counselor Trainees	61.82	9	6.87	4.33	.004
error	28.55	18	1.59		
Total	103.84	29			

Figure 1: Mean (\bar{X}), Range (R), and Standard Deviation (SD) of Percentage Differences Between Actual Performance and the Standard and ANOVA of Total Group at Each Observation Time

The change is also educationally significant because it lends support to a particular approach to counselor training. Although the counseling behavior of trainees as a group changes toward the standard between the first observation and the observation after training, the difference is not

statistically significant. This conflicts with the initial hypothesis. What seems to have happened (see Figure 1) is that while the mean for the group did not change significantly, those trainees most unlike the model in the beginning (C, E, F, J, and I) became more like the model after training. Those trainees most like the model in the beginning (A, G) have changed moderately away from the model (see Table 2).

Regression toward the mean was considered as a possible explanation for the change. However, this explanation was rejected because the following conditions for such an argument did not seem to be present in the study: (1) regression frequently occurs when "subjects are selected because they deviate from the mean" (Hopkins and Glass, 1978: 167); (2) regression is most typical of pretest-posttest research designs (Hopkins and Glass, 1978); (3) regression is typically "a negatively accelerated function of elapsed time" (Campbell and Stanley, 1963: 41). The fluctuation in means between observations one, two, and three, along with the effects of treatment being greater at observation three than at one or two, seem to further discount the regression toward the mean explanation.

In addition to these findings, it was discovered that trainees actually performed less like the standard after some experience, but before microcounseling training. This finding conflicts with popular notions that supervised experience alone is sufficient for trainees to improve the quality of their performance, and supports the belief that single skill training is necessary.

Counselor training program evaluation suffers from problems similar to all experimental field studies utilizing small groups of human subjects. Simple pretest-posttest designs or control group designs often have not been satisfactory approaches to this problem. One solution is to observe the same group of trainees over time and determine the impact of microcounseling training by comparing their behavior to a predetermined standard of behavior. This approach allows the trainer to compare the trainee group to itself at several times during training and to conduct an objective assessment of the trainee's ability to perform counseling skills.

For purposes of this investigation, a standard for counseling has been defined in terms of the frequency of selected microcounseling skills. One could argue that the specific frequencies should be different, or even that a skill frequency model cannot include all of the important elements of effective counseling. This study does not attempt to respond to either of these issues. Instead, the authors welcome debate around the evaluation model that is utilized in this study.

REFERENCES

- AUTHIER, J., and K. GUSTAFSON (1976) "The application of supervised and nonsupervised microcounseling paradigms in the training of registered and licensed practical nurses." *J. of Consulting and Clinical Psychology* 44: 704-709.
- CAMPBELL, D. T. and J. C. STANLEY (1963) *Experimental and Quasi-Experimental Designs for Research*. Chicago: Rand McNally.
- COOK, T. D., and D. T. CAMPBELL (1979) *Quasi-Experimentation: Design and Analysis Issues for Field Settings*. Chicago: Rand McNally.
- FERGUSON, G. A. (1976) *Statistical Analysis in Psychology and Education*. New York: McGraw-Hill.
- HAASE, R. and O. DiMATTIA (1970) "The application of the microcounseling paradigm to the training of support personnel in counseling." *Counselor Education and Supervision* 10: 16-22.
- HARRIS, R. H. (1975) *Primer of Multivariate Statistics*. New York: Academic.
- HERSEN, M. and D. H. BARLOW (1976) *Single Case Experimental Designs*. New York: Pergamon.
- HOPKINS, K. D. and G. V GLASS (1978) *Basic Statistics for the Behavioral Sciences*. Englewood Cliffs, NJ: Prentice-Hall.
- IVEY, A. E. and J. AUTHIER [eds.] (1978) *Microcounseling*. Springfield, IL: Charles C Thomas.
- IVEY, A. E., C NORMINGTON, C. MILLER, W. MORRILL and R. HAASE (1968) "Microcounseling and attending behavior: an approach to prepracticum counselor training." *J. of Counseling Psychology* 15, 2: 1-12.
- KASDORF, I. and K. GUSTAFSON (1978) "Research related to microtraining," in A. E. Ivey and J. Authier (eds.) *Microcounseling*. Springfield, IL: Charles C. Thomas.
- KERREBROCK, R. (1971) "Application of the microcounseling method using videotape recordings to the training of teachers in basic counseling techniques." Ph.D dissertation, University of Southern California. *Dissertation Abstracts International*, 1971, 32, 740A. (University Microfilms No. 71-21, 470).
- MARASCUILO, L. (1971) *Statistical Methods for Behavioral Science Research*. New York: McGraw-Hill.
- MORELAND, J., J. PHILLIPS, A. IVEY and J. LOCKHART (1970) "A study of the microtraining paradigm with beginning clinical psychologists." University of Massachusetts. (unpublished)
- NUTTALL, E. V. and A. E. IVEY (1978) "Research for action: the tradition and its implementation," in L. Goldman (ed.) *Research Methods for Counselors*. New York: John Wiley.
- SCROGGINS, W. and A. IVEY (1976) "An evaluation of microcounseling as a model to train resident staff." University of Alabama. (unpublished)

Stephen Joel Gill is an Assistant Professor in the School of Education at the University of Michigan. His interests are in college student services, teaching helping skills, and program evaluation.

Carl F. Berger is a Professor in the School of Education at the University of Michigan. His interests are in research design, science education, and the application of microcomputers to teaching.

George L. Cogar is a Psychologist with the Veterans Administration and is in private practice in Tomah, Wisconsin.