ABSTRACT
              The paper outlines several studies over a two-year
period to develop a self-report and observer-rating measure of
sensitivity/encounter group outcome. The initial form of the scale
was taken from McMillan (1971) who developed a measure of 16
categories of group outcome; McMillan's work indicated the scale had
high reliability. Subsequent study supported somewhat the validity of
the scale, but also indicated several areas where the scale could be
improved. The scale was shortened and given a checklist format;
validity was again supported in further studies. Finally, the scale
was converted to behavioral items, submitted to a large sample for
item and factor analysis calculations. This form yielded high
reliability, and 11 factors from an image analysis solution.
Extensive forthcoming research to further investigate the reliability
and validity of the scale was outlined. It was concluded that the
scale assessed a number of small factors relating to effective
functioning in interpersonal relationships which facilitate intimacy.
(Author)

Development of a Behavioral Affective Relationship Scale
for Encounter Research

William R. Shadish, Jr., and Thomas Zarle
Purdue University

The present paper will outline a program of research consisting of several

studies over the last two years that have had the goal of developing a

self-report and observer-rating measure of sensitivity/encounter group outcome.

It was specified that the measure should be clearly related to group goals

as stated in the literature, meet or exceed accepted standards of reliability

and validity, and be constructed with sound statistical and methodological

techniques. This measure might also be useful in such related fields of

research as group therapy, skill training groups, task and discussion groups,

and any fields in which an assessment of human affective relations would be

needed.

The first task was specification of group goals as stated in the

literature; this has been accomplished previously by McMillan (1971) who

completed an extensive review of the group literature and synthesized 24

goal categories. He also devised the McMillan Affective Relationship Scale

(MARS) which consisted of 96 attitudinal and behavioral items representative

of the goals. Finally, McMillan demonstrated that the MARS was reliable over

time (median $r = .92$), showed the 24 categories to be internally homogeneous,

and showed that the MARS could measure change in subjects participating in

encounter groups.

Three major studies between 1974 and the present time have been conducted

by the present authors to replicate and extend McMillan's work. These studies

2

will be reviewed briefly and then research being conducted between now and next year will be outlined. However, the important contribution of McMillan's initial work must be acknowledged, as he provided a solid and adequate data source from which to create a measure of group outcome.

In the Fall of 1974, a methodologically rigorous study (Shadish, 1975) further supported the validity of the MARS. Twenty-four male and twenty-five female undergraduate volunteers were assigned randomly to one of six 10-hour structured verbal or non-verbal encounter groups. Two extensively trained leaders were assigned randomly to each group. The results were briefly as follows. First, concurrent validation of the scale against the Eysenck Personality Inventory revealed a pattern of responses consistent with previous research concerning anxiety and group training (Zarle, 1975). There was no significant correlation between the MARS and the Eysenck Lie Scale. The MARS failed to distinguish between verbal and non-verbal group members, replicating the results of McMillan; but it did distinguish benefits for encounter partici-pants compared to delayed-treatment controls. No effect for pretesting seemed present; that is to say that use of the Solomon four-group design indicated that pretesting did not seem to influence subsequent administrations of the MARS. The following difficulties with the MARS were also noted. A form constructed for observer rating proved unreliable over raters; and the scale seemed to foster a positive response bias.

In the Fall of 1975, 28 members were assigned randomly to one of 8-hour encounter groups. These groups were not structured by the experimenter and were conducted in marathon format. Two groups spontaneously utilized exercises more than 25% of the time, while the other two used almost no exercises; this difference did not produce a statistically significant effect.

3

The form of the MARS used in this study differed from previous forms as follows. First, a checklist format was used in place of the original 9-point Likert-type format, in hopes of reducing the positive response bias. Second, the number of items was reduced to 23 since a pilot study factor analysis indicated the possible presence of one general factor; each item thus represented one goal category. The results of this study were as follows. Members of the groups showed no change in total score from pretest one (two weeks prior to the group) to pretest two (immediately prior to the group); their scores increased from pretest two to immediate posttest, and this change was maintained at two week posttest. Additionally, the scale correlated significantly with the Self-Testimony scale, which was the best predictor of outcome in the encounter group study of Lieberman, Yalom, & Miles (1973). There was no signifi.. correlation between the MARS and either the Eysenck Lie Scale or the Marlowe-Crowne Social Desirability Scale.

Because the previous two studies seemed to support the relevance of McMillan's goal categories to group research and also seemed to support further development of a reliable measure, a study was designed to clarify the factor structure and internal reliability of a behavioral scale derived from McMillan's goal categories. In December of 1975, 14 encounter group leaders responded to a questionnaire instructing them to list one observable or audible behavior that a group member could do to illustrate each goal. There was much subjective agreement among leaders as to the type of behavior which would illustrate each category. Two such behaviors were selected for most categories; for a few goal categories only one behavior could be agreed upon. In January of 1976, more than 800 Introductory Psychology students responded on a five point format ranging from "Would never behave that way"

to "would always behave that way, with no exceptions." 786 subjects were
retained after incomplete scales were discarded. Internal consistency
reliability analyses yielded a Cronbach's alpha of .92 and a Guttman lower
bound reliability of .94. Second, the 41 item scale was submitted to a
variety of factor analytic techniques. Suggestions by Gorsuch (1972)
concerning analytic techniques and number of factors were followed; and
according to a wide variety of criteria, the best solution was an image
analysis extracting 12 factors and subjected to an oblique rotation. These
factors have been tentatively identified as follows:

1. Expression of positive feelings
2. Effective communication to others
3. Discussing others feelings
4. Closeness expressed to others
5. Acceptance and liking for self
6. Understanding what happens between others
7. Expression of negative feelings
8. Expression of change in attitude
9. Requesting critical feedback
10. Discussing self and relationship to others
11. Expressing risky feelings
12. Acceptance of criticism and disagreement

This last factor seemed to be the result of a fortuitous phrasing of two items
and was dropped from further consideration. Second order factor analyses did
not lead to the emergence of consistent factors. However, in all cases, the
first factor accounted for 50% to 75% of the unrotated variance, supporting
an interpretation of a single underlying second order factor. Further evidence
for this interpretation was the nearly zero determinant (8×10^{-7}) of the
first order correlation matrix, the high internal consistency reliability
coefficients, and high correlations among second order factors. The first
order factors seemed to adequately represent McMillan's meaning in his goal
categories, and these were retained for further scale development.

5

This completes a review of past major research in the development of an affective relationship scale based on McMillan's goal categories. A fourth study is being conducted at the present time involving the use of confirmatory factor analytic techniques to test the hypothesis that the factors cited above are sufficient to account for the variance in the scale. Furthermore, the number of items per factor will be equalized at four to six items per factor, yielding a 44 to 66 item scale. The internal consistency reliability of the scale total score is expected to be high; however, reliability estimates for each factor will probably be less impressive due to the small number of items per factor. This was deemed necessary in order to keep the scale at a reasonable length. During the forthcoming academic year, a series of scaling analyses and reliability and validity studies will be conducted using the new scale.

Having completed a review of past and proposed research with the scale, it seems appropriate to make a few comments about what the scale seems to measure. The items and the response format most closely resemble a behavioral frequency count. Thus the responder indicates how often he/she does a certain behavior. These seem to represent behaviors defined by group leaders as desirable interpersonal relationship behaviors. In general, then, the scale seems to measure a number of smaller factors related to effective functioning in interpersonal relationships.

The above discussion will hopefully give a sense of the amount and quality of research that has been completed and is currently being done in developing a scale that is useful for group outcome research. Detailed descriptions of the research cited in this paper are in preparation and are available from the authors.

REFERENCES

Gorsuch, R. Factor Analysis. Philadelphia, Pennsylvania: W.B. Saunders,
    1974.

Lieberman, M.; Yalom, I., and Miles, M. Encounter Groups: First Facts.
    N.Y.: Basic Books, Inc., 1973.

McMillan, O.D. "The development and use of the McMillan Affective Relationship
    Scale in measuring the effects of verbal interaction and of selected
    non-verbal techniques of communication on synthesized desirable group
    outcomes of group dynamics procedures in sensitivity training." East
    Texas State University, Ph.D., Dissertation, 1971. Dissertation Abstracts
    International, Order No. 72-10, 802.

Shadish, W.R. "The Validity of the McMillan Affective Relationship Scale in
    Encounter Group Outcome Research." Unpublished Masters Thesis, Purdue
    University, 1975.

Zarle, T. and Willis, S. Induced affect as a pretraining technique for
    encounter group stress. Journal of Counseling Psychology, 1975, 22, 49-53.