

COUNSELLOR COMMUNICATION STYLE AS A DETERMINANT OF RATER-PERCEIVED EMPATHY

DAVID L. RENNIE, HARLEY L. BURKE and SHAKÉ G. TOUKMANIAN

York University

Abstract

Carkhuff-scale empathy of 64 counsellors was judged by two randomly sampled raters from each of two sources of raters. Each counsellor was evaluated both with his client present in the transaction (unedited tape material) and with his client's utterances erased (edited material). In a second phase of the study, the average empathy scores of the four raters were regressed against three linguistic and three paralinguistic attributes of counsellor communication. Separate regressions were conducted for the empathy ratings given to the unedited versus edited material. The results of the first part of the study paralleled earlier findings of Truax indicating that when empathy ratings on the unedited material were correlated with the ratings in the edited material, the result approximated the interrater reliability of the scale. The results of the two multiple regressions were similar and supported the hypothesis that, in rating unedited material, empathy raters focus more on the style of the counsellor's response than on the relationship between the response and the client utterance preceding it. The validity of this deductive method of discriminating is discussed.

Résumé

En utilisant l'échelle de l'empathie de Carkhuff, deux évaluateurs choisis au hasard parmi deux groupes d'évaluateurs, ont évalué 64 conseillers. Le niveau d'empathie des conseillers fut évalué à l'aide d'enregistrements sur bande sonore. Certaines bandes contenaient les échanges du client (bande nonmodifiée). Sur les autres bandes, on avait effacé les interventions des clients (bande modifiée). Dans la seconde phase de l'étude, on a soumis séparément à une analyse de régression multiple en fonction de trois variables de communication linguistiques et de trois variables paralinguistiques les valeurs moyennes obtenues par les évaluateurs tant sur les bandes modifiées que nonmodifiées. Les données recueillies dans la première phase de l'étude étaient comparables aux résultats de Truax. Elles indiquaient que lorsque des évaluations de l'empathie faites sur des échanges nonmodifiés sont comparées aux évaluations faites sur des échanges modifiés, le résultat se rapproche de la valeur "interrater reliability" de l'échelle. Les deux analyses de régression multiple produisaient des données semblables appuyant ainsi l'hypothèse que les évaluateurs de l'empathie qui utilisent des échanges verbaux nonmodifiés, basent leurs jugements davantage sur le style de communication des conseillers que sur le rapport qui existe entre la réponse des conseillers et celle de son client. Enfin, on discute la validité de cette méthode déductive de discernement.

Accurate empathy has been defined as a phenomenon that "... involves both the therapist's *sensitivity to current feelings* and his *verbal facility to communicate this understanding in a language attuned to the client's current feelings*" (Truax & Carkhuff, 1967, p. 46). This definition would seem to make it essential that if an external rater is to evaluate therapist empathy, the locus of evaluation must reside in the communicational transaction between the client and his therapist. However, Truax (1966) found that when four raters using his accurate empathy scale (Truax, 1961) evaluated segments of client-therapist transactions (unedited material) and the same segments with the client utterances erased

(edited material), the correlation between the empathy ratings on the two sets of material was .68, which approximated the reliability of the scale. Chinsky and Rappaport (1970) and Rappaport and Chinsky (1972) claim that this finding constituted evidence that Truax's (1961) raters were responding to some aspect of counsellor communication, such as communicational style, that was not implied by the definition of empathy underlying the scale.

There is growing evidence that counsellor communication style is an important set of cues to which empathy raters attend (Toukmanian & Rennie, 1975; Uhlemann, Lea, & Stone, 1976; Wenegrat, 1974; Zimmer & Anderson, 1968).

However, these studies used unedited material as the stimulus complex presented to the raters. This approach confounds the counsellor communication style with the effect of that style on client responding. Truax's (1966) study is the only reported investigation that eliminated this confounding. However, there are a number of methodological features of his study that weaken both the internal and external validity (Campbell & Stanley, 1966) of his findings.

With respect to internal validity, his use of just five therapists as the source of his 50 unedited and 50 edited segments may have enabled his raters to recognize the voice of each therapist and hence to stereotype each therapist as "good" or "bad" in the unedited transactions. This stereotyping may in turn have generalized to the perceptions of the therapists in the edited material thereby resulting in a spurious estimate of the association between the ratings on the two sets of stimuli. Second, Truax (1966) counterbalanced the order of presentation of the set of unedited segments and the set of edited segments. This approach appears commendable but it is possible that the raters who heard the edited material first could have developed a set to focus on only the therapists when they rated unedited segments. Whether or not this method created a spurious effect is unclear in view of Truax's (1966) failure to analyze for the effect of order of presentation. Finally, Truax's (1966) editing procedure was such that, while the client responses were eliminated, the ratio of therapist talk-time to the total time of the interview segment was left intact vis-à-vis the unedited tape. While the relationship between therapist activity and rated empathy is not altogether clear (Caracena & Vicory, 1969; Truax, 1970; Hargrove, 1974; Wenegrat, 1974) this uncontrolled variable could have contributed to Truax's (1966) results.

In terms of external validity, there is no evidence that Truax (1966) randomly selected his four raters from the population of raters, hence it is difficult to make any conclusions about the generalizability of his results.

The present study was conducted in two parts. In Part 1, the objective was to demonstrate that, under conditions of better experimental control, results similar to those of Truax (1966) could be obtained with the Carkhuff Empathic Understanding Scale (Carkhuff, 1969). Given positive findings in Part 1, Part 2 was designed to examine the basis for the tendency of empathy raters to yield highly correlated ratings on unedited versus edited material. It was speculated that, when rating unedited material, raters tend to respond primarily to the style of a counsellor utterance rather than to the relationship between it and the client response preceding it. This being the case, the same set of salient cues is available to the

raters when they listen to edited material, which in turn produces the high correspondence between ratings on the two types of material. To test this proposition, empathy rated on unedited material was regressed against three linguistic and three paralinguistic attributes of counsellor communication and a similar analysis was conducted for empathy rated on edited material. It was hypothesized that the pattern of correlations between the six predictors and empathy rated on unedited segments would be similar to the pattern of association between the predictors and empathy rated on edited material.

METHOD

Part 1

Experimental Design

Part 1 of the study has a 2 (Source of raters) × 2 (Raters nested within sources) × 2 (Unedited-edited) × 2 (Same different counsellor utterances) design. Each of the four factors is described below.

The Stimulus Material

The tape segments were derived from the 7-minute posttest interview segment of each of 64 of the counsellors serving as subjects in an earlier study (Rennie & Toukmanian, Note 1). Forty-four of the counsellors were fourth year arts students enrolled in an undergraduate course in counselling: 10 were graduate students engaged in a clinical-counselling practicum at the Counselling and Development Centre of York University; and 10 were non-psychology majors enrolled in an undergraduate course in urban organization. These counsellors had been randomly distributed across nine undergraduate students who volunteered to serve as clients and who were paid for their work. Each client presented a problem which was his unique expression of the theme "interpersonal difficulty with someone significant in my life" that was common to all standard clients.

Procedure

Editing. These counselling transactions had been recorded on monaural magnetic tape. To erase client utterances, a machine playing the unedited tape was "patched" into a second recorder and only the counsellor utterances were rerecorded. Through careful work the editor eliminated all client utterances except those that directly overlapped counsellor talk. In order to control for the ratio of counsellor talk-time to total interview time, each rateable counsellor utterance (i.e., a counsellor response occurring between two client utterances) was separated by a standard 3 to 5 second interval.

The interviews of the 64 counsellors were

randomly assigned to two groups of 32. For each member of one group, one of the two halves of the 7-minute interview was randomly selected and this segment was edited. Hence, for this group, both the unedited and edited material came from the same half of the 7-minute interview. For each member of the other group, one of the two halves of his 7-minute interview was randomly selected and the other half of his 7-minute interview was edited. Thus, for this group, the randomly selected half of the interview was used as the unedited segment, and the other half of the 7-minute interview was edited. The editing *per se* explains the unedited-edited factor in the design. The differential use of the two halves of the 7-minute interview explains the same-different counsellor utterance factor.

The latter factor was used as a control against the possibility that, despite the interference produced by listening to many different counsellors, the empathy raters might remember the rating given to a counsellor on an unedited segment and thereby bias his estimate of empathy heard on the corresponding edited segment. It was assumed that "true" empathy would not appreciably vary over a 7-minute interview.

Each unedited and edited segment was identifiable by a numerical code. As a further control against rater memory effect, the sequential order of segments on the edited master tape differed from the order of segments on the unedited master tape.

Empathy raters. The Guelph Human Services Community under the direction of Dr. Ralph Bierman and the contactable population of York University graduate students who had taken a Carkhuff-based human relations training programme were selected as two sources of raters. The group had didactic training in rating empathy with the Bierman (Note 2) revision of the Carkhuff Empathic Understanding Scale. The members of the York group were trained to use the Carkhuff Scale during their training programme. A list of 12 potential raters were secured from an official of the Guelph organization which compared with the list of 15 members of the York group. The members of each of these lists were randomized and the experimenter worked down each randomized list until he found two raters from each group who were willing to do the task. This procedure explains the 2 (Source of raters) and 2 (Raters nested within sources) factors in the design.

Each rater was given the unedited master tape together with instructions to independently rate the material using an enclosed copy of the Carkhuff Scale. The edited mastertape was withheld until the unedited mastertape was returned together with the ratings of it. As an

additional precaution, the raters were admonished to destroy any copies of their ratings of the unedited tape before proceeding with the edited master.

Part 2

Instrumentation

Linguistic measures. The Ivey-based taxonomy of counsellor utterances developed by Toukmanian and Rennie (1975) was used to assess linguistic aspects of counsellor communication. The taxonomy consists of open invitation to talk, closed questions, interpretation and advice, and other. Toukmanian and Rennie (1975) found 75.2% interrater agreement and interrater reliabilities between .75 and .89 for the four category system.

Paralinguistic measures. Paralinguistic features of counsellor communication were assessed using the Therapist Voice Quality classification system (Rice, 1965; 1973). This 9-category system pertains to qualities of communication such as energy and pitch independent of the lexical content of the communication. For analysis, eight categories are collapsed to Expressive Voice, Usual Voice, and Distorted Voice, with the ninth category serving as an "other" classification. Rice (1965) found that raters could be trained to categorize at a level wherein, at a minimum, their agreements were significantly above chance expectancy.

Procedure. Two trained undergraduates independently used the Toukmanian-Rennie taxonomy. Since previous research (Toukmanian & Rennie, 1975) had indicated high interrater agreement and reliability for these categories, a rerating of the edited segments was judged unnecessary.

Two graduate students, trained in the use of the Rice Therapist Voice Quality Scale, used a key-worded transcript to independently apply the scale to each 3.5 — minute segment. Since the investigators were less experienced in the use of this scale, these raters rated both the unedited and edited segments for each of the 64 counsellors.

RESULTS

Part 1: Ratings of Empathy

Interrater reliabilities. The intraclass correlations (Ebel, 1951) between the empathy scores of the four raters, taken in pairs, for the unedited segments ranged from .19 to .50 ($\bar{X} = .32$) and from .16 to .53 ($\bar{X} = .33$) for the edited material. When the intraclass technique was used to determine the joint reliability of the four raters taken simultaneously, the resulting correlations were .65 and .69 for the unedited and edited material, respectively. This latter approach appears to be the one used in the Truax (1966)

study wherein Truax obtained corresponding values of .76 and .66.

Components of variance in empathy judgments. Table 1 shows the means and standard deviations of the empathy ratings partitioned into Source of raters, Raters nested within sourced, Edited-unedited material and Same-different counsellor utterances. It can be seen that for all material, the Guelph raters yielded an average empathy score of 1.67 (S.D. = 0.62) while the average score for the York raters was 2.35 (S.D. = 0.82).

Table 1
Means and Standard Deviations of Empathy Ratings

Group	Editing	Raters nested within Sources							
		Guelph				York			
		Rater 1		Rater 2		Rater 3		Rater 4	
	M	SD	M	SD	M	SD	M	SD	
"Same" Segments (n = 32)	Unedited	1.63	.57	1.65	.50	2.45	1.05	2.32	.79
	Edited	1.50	.53	1.61	.68	2.42	.83	2.24	.69
"Different" Segments (n = 32)	Unedited	1.90	.70	1.78	.80	2.25	.84	2.45	.85
	Edited	1.62	.52	1.66	.69	2.28	.81	2.42	.71

An analysis of variance of the four factor design revealed a single main effect for Source of raters $F(1,62) = 156.5$; $p < .01$. Calculation of ω^2 (omega squared) using Kirk's (1968) formula showed that this factor accounted for 97.5% of the variance contributed by the four factors of the design.

Relationship between raters' judgments on the unedited and edited materials. The average score of the four raters for each unedited and edited tape segment was determined. The mean empathy score of each rater's unedited segment was correlated with the mean empathy rating of his edited segment. The correlation between the "Same" unedited and edited segments was $r = .61$, $p < .001$ (two-tailed) which compares favourably with Truax's value of .68. The correlation between the "Different" unedited and edited segments was $r = .51$; $p < .01$, (2-tailed).

Part 2: Linguistic and Paralinguistic Correlates of Empathy

Interrater reliabilities. For the unedited segments, the percentage of rater agreement of 478 counsellor responses classified into open invitation to talk, closed questions, interpretation and advice, and other categories was 86.6%. Interrater reliability was also examined using, for each subject, the total frequency in each of the first three categories ("other" was excluded because of few entries). The intraclass correlations for open invitation to talk, closed questions and interpreta-

tion and advice were .92, .92, and .75, respectively. For the edited segments the corresponding values were 86.3% (491 counsellor responses) and $r = .83$, .94, and .66.

With respect to therapist voice quality, the raters showed 72.3% agreement on 419 responses in the unedited material classified into expressive, usual and distorted categories and 67.6% agreement on the categorization of 463 responses in the edited material. The interrater reliabilities, calculated with the same method as that used for the linguistic predictors, were .86 (expressive), .70 (usual) and .84 (distorted) for the unedited segments and .83, .75 and .91 respectively, for the edited material.

Regression of empathy against the six predictors. In determining the extent to which the four raters in the present study used counsellor cues similarly for both the unedited and edited material, the average ratings of the four empathy raters were regressed against the three linguistic and the three paralinguistic categories. A separate step-wise multiple regression was made on empathy ratings given to the unedited material vis-à-vis the ratings given to the edited segments. A summary of the step-wise regressions is shown in Table 2. It can be seen that the predictors yielded a R of .56, $F(6,57) = 4.41$; $p < .001$ for the unedited material and an R of .56, $F(6,57) = 4.45$; $p < .001$ for the edited material. Open invitation to talk and usual voice emerged as the first and second strongest predictors in both analyses. These two predictors contributed 65% and 93% of the accountable variance in the analysis of the unedited material and edited material, respectively.

The array of Pearson correlation coefficients signifying the relationship between each predictor and empathy reveals that the discrepancy between a given coefficient for unedited material vis-à-vis the coefficient for edited material ranged from .01 for open invitation to talk to .19 for interpretation and advice. A Fisher r to z transformation indicated that the probability associated with the latter discrepancy was $z = 1.50$; $p < .07$, (2-tailed).

Inspection of the pattern of partial correlation coefficients indicates that the discrepancy between the members of the pairs of coefficients ranged from .04 for closed questions to .20 for usual voice. The two-tailed probability associated with the latter discrepancy was $z = 1.58$; $p < .06$.

Hence, the step-wise multiple regression analyses and the supplemental pattern of Pearson and of partial correlations for the unedited and edited material support the hypothesis that similar cues would be used by the empathy raters whether or not the client was present in the transaction.

Table 2

Relationship between Six Predictors and Empathy Rated
on Unedited and Edited Material^a

Predictor	Multiple R	Rank of Predictor	Pearson r	Partial r	Beta
Open invitation to talk	.39 (.38)	1 (1)	.39 (.38)	.22 (.37)	.26 (.44)
Usual voice	.45 (.54)	2 (2)	-.07 (-.21)	-.01 (-.21)	-.01 (-.27)
Distorted voice	.49 (.56)	3 (6)	-.26 (-.08)	-.02 (.09)	-.02 (.09)
Interpretation and advice	.53 (.56)	4 (5)	-.24 (-.05)	-.29 (-.10)	-.27 (-.09)
Expressive voice	.53 (.56)	5 (4)	.32 (.34)	.21 (.13)	.28 (.15)
Closed questions	.56 (.55)	6 (3)	-.21 (-.26)	-.19 (-.15)	-.25 (-.17)

^aThe values for edited material appear in parentheses.

* $p < .05$

** $p < .01$

DISCUSSION

The results of Part 1 paralleled those of Truax (1966) under conditions which provided relatively better control over spurious sources of correlations between ratings on unedited and edited material. Furthermore, our results emerged from the use of raters who were randomly selected from two sources, which strengthens the generalizability of the results somewhat, although it would have been increased further had we randomly sampled from the population of sources.

A cautionary note is that, while the joint interrater reliability of our four raters approximated Truax's (1966) findings, the pair-wise reliability of our (and, we presume, Truax's) raters were low relative to those typically reported in the literature (e.g., Berenson, Carkhuff & Myrus, 1966; Fry, 1973). Further work is required before the results of the two studies can be safely generalized to raters showing high interrater reliability.

The results of Part 2 of the study appear to shed some light on the cues empathy raters use when making their judgments. Turning to the three linguistic attributes first, the finding that open invitation to talk correlated positively and closed questions and interpretation and advice correlated negatively can be understood in both lexical and metacommunicational terms. From the lexical standpoint, a rejection of feeling has the formal appearance of being empathic, since empathy is defined in terms of the client's feelings. Similarly, to the extent that open-ended questions and bids for clarification have the client's feelings as part of their lexical content, the rater is in a position to assume that counsellors using these types of leads

are focusing on what they are supposed to focus on in their attempts to be empathic.

However, closed questions and interpretation and advice can also have the client's feelings as lexical content, yet these categories correlate negatively with empathy. This finding appears to be explicable by the differing metacommunicational messages imparted by open invitation to talk versus the other two categories. Open invitation to talk responses stimulate the client to engage in divergent exploration whereas the other two categories stimulate him to converge his thinking. Putting it in terms of client centred theory, the divergence-producing responses stimulate the client to expand his frame of reference while the convergence-producing responses limit it. Interpretation and advice, in addition to stimulating convergence, force the client to accommodate his self-awareness to the counsellor's own frame of reference, which may help explain why even low frequencies of this type of counsellor utterance can dramatically affect his or her rated empathy score.

The relevance of the three paralinguistic attributes appears relatively more straightforward. Expressive voice, which depicts an energized, caring voice, correlates positively with empathy. Conversely, distorted voice which is characterized by thinness, low-energy and unusual pitch levels, correlates negatively.

Zimmer and Anderson (1968) found that counsellor communication style variables similar to our linguistic attributes appeared to operationize empathy. However, Hargrove (1974) found that what appears to be a reasonably strong association between rated empathy and the

counsellor's silence, length of response, and frequency of interrupting. It should be emphasized that our six correlates of rater-perceived empathy merely sampled the universe of attributes and that the discovery of this domain awaits further research.

The second major finding of Part 2 was that our raters apparently used these six attributes to essentially the same extent whether or not they heard the client in the transaction. We thus have evidence that our raters were judging attributes of counsellor communication style more so than empathy as theoretically defined and this evidence supports the challenge of the validity of rater-perceived empathy as posed by Chinsky and Rappaport (1970) and other (e.g., Kurtz & Grummon, 1972; Barret-Lennard, Note 3).

It appears that when rating the unedited material, our raters did not inductively match the accuracy of each counsellor response with the preceding client statement(s) serving as a referent. Instead, the raters seem to have had a model of the stylistic components of empathic responses and to have deduced the level of counsellor empathy on the basis of the extent to which the rater's style matched that prescribed by the model.

Can the approach to rating empathy conceivably produce accurate ratings? Logically, this method could produce high accuracy only if the components of counsellor style are independent of client utterances. Clearly, the linguistic components are not independent in this way. For example, a reflection of feeling may have the stylistic indices of a highly empathic response but reflect the wrong feeling. The same is true of counsellor voice tone. Voice tonal expressions of interest and caring may contribute to empathy but, in order for these attributes to be judged empathic, they must be associated with the linguistic attributes of the utterance to determine their accuracy. Failing this, the interest and caring would be legitimate indicators of other counsellor attributes such as warmth or respect, but not empathy.

Hence, highly accurate judgments on empathy on the basis of counsellor style necessitate the presence of client referent statements. However, probabilistic estimates of empathy could be made on the basis of style in the absence of such referents. It is plausible that should counsellors use a style that encourages client divergent exploration, the chances are greater than they are being empathic than if they use a style that forces convergent thinking. Similarly, while voice tone associated with linguistically inaccurate empathy cannot be considered a component of empathy, it is also true that voice tone modifies linguistically accurate empathy. Hence, the empathy rater will likely be influenced in the right direction if he focuses on style to the exclusion of the interaction

between the counsellor and client. However, for the rater to do so constitutes task simplification since this deductive approach prevents the rater from attending to the nuances of the transaction that makes up the subject matter of accurate empathy as theoretically defined.

In this context, Truax (1972) is somewhat perplexing when he maintains that, in training empathy raters, he trained them to not listen to the client but only to the counsellor. This training procedure abrogates the theoretically anchored dictates of the scale. In our view, empathy raters should be encouraged to take on the tough intellectual exercise of rating empathy inductively since only then can the scale accurately reflect the essence of the counselling transaction.

Reference Notes

1. Rennie, D.L., & Toukmanian, S.G. *Acquisition of counselling skills by introverted and extraverted counsellor trainees under conditions of experiential and didactic training*. Paper presented at the annual meeting of the Canadian Psychological Association, Windsor, Ontario, June 1974.
2. Bierman, R. *Personal interaction manual*. Unpublished report, University of Waterloo, 1971.
3. Barrett-Lennard, G.T. *Empathy in human relationships: Significance, nature and measurement*. Paper presented to the annual conference of the Australian Psychological Association, Perth, August 1974.

References

- Berenson, B.G., Carkhuff, R.R., & Myrus, P. The interpersonal functioning and training of college students. *Journal of Counseling Psychology*, 1966, 13, 441-446.
- Campbell, D.T., & Stanley, J.C. *Experimental and quasi-experimental designs for research*. Chicago: Rand McNally, 1966.
- Caracena, P.F., & Vicory, J.R. Correlates of phenomenological and judged empathy. *Journal of Counseling Psychology*, 1969, 16, 510-515.
- Carkhuff, R.R. *Helping and human relations (Vol. 1)*. New York: Holt, Rinehart & Winston, 1969.
- Chinsky, J.M., & Rappaport, J. Brief critique of the meaning and reliability of "accurate empathy" ratings. *Psychological Bulletin*, 1970, 73, 379-383.
- Ebel, R.L. Estimation of the reliability of ratings. *Psychometrika*, 1951, 16, 407-424.
- Fry, P.S. Effects of desensitization treatment on core condition training. *Journal of Counseling Psychology*, 1973, 20, 214-219.
- Hargrove, D.S. Verbal interaction analysis of empathic and nonempathic responses by therapists. *Journal of Consulting and Clinical Psychology*, 1974, 42, 305.
- Kirk, R. *Experimental design: Procedures for the behavior sciences*. Belmont, California. Brooks/Cole, 1968.
- Kurtz, R.R., & Grummon, D.L. Different approaches to the measurement of therapist empathy and their relationship to therapy outcomes. *Journal of Consulting & Clinical Psychology*, 1972, 39, 106-115.

- Rappaport, J., & Chinsky, J.M. Accurate empathy: Confusion of a construct. *Psychological Bulletin*, 1972, 77, 400-404.
- Rice, L.N. Therapist's style of participation and case outcome. *Journal of Consulting Psychology*, 1965, 29, 155-160.
- Rice, L.N. Client Behavior as a function of therapist style and client resources. *Journal of Counseling Psychology*, 1973, 20, 306-311.
- Toukmanian, S.G., & Rennie, D.L. Microcounseling versus human relations training: Relative effectiveness with undergraduate trainees. *Journal of Counseling Psychology*, 1975, 22, 345-352.
- Truax, C.B. A scale for the measurement of accurate empathy. *Psychiatric Institute Bulletin*, Wisconsin Psychiatric Institute, University of Wisconsin, 1961, 1, No. 12.
- Truax, C.B. Influence of patient statements on judgments of therapist statements during psychotherapy. *Journal of Clinical Psychology*, 1966, 22, 335-337.
- Truax, C.B. Length of therapist response, accurate empathy and patient improvement. *Journal of Clinical Psychology*, 1970, 26, 539-541.
- Truax, C.B. The meaning and reliability of accurate empathy ratings: A rejoinder. *Psychological Bulletin*, 1972, 77, 397-399.
- Truax, C.B., & Carkhuff, R.R. *Toward effective counseling and psychotherapy*. Chicago: Aldine, 1967.
- Uhlemann, M.R., Lea, G.W., & Stone, G.L. Effects of instructions and modeling on low-functioning trainees. *Journal of Counseling Psychology*, 1976, 23, 509-513.
- Wenegrat, A. A factor analytic study of the Truax Accurate Empathy Scale. *Psychotherapy, Theory, Research and Practice*, 1974, 11, 48-51.
- Zimmer, J.M., & Anderson, S. Dimensions of positive regard and empathy. *Journal of Counseling Psychology*, 1968, 15, 417-426.

Footnotes

This paper is based on a thesis submitted by H. Burke in partial fulfillment of the requirements of the master's degree of York University. The study was supported by Canada Council grants Nos. S73-1809 and S74-1820 awarded to D. Rennie and S. Toukmanian. The authors extend their thanks to Lance Mitson for his assistance in the data analyses.

1. Truax (1966) is unclear about the meaning of the term "reliability of the scale." While he trained his raters to a rate — rerate reliability of .50 or greater, it appears that he compared interrater reliability to the correlation between the ratings on the unedited versus edited material.