Reflective Listening in Counseling: Effects of Training Time and Evaluator Social Skills

ERIK RAUTALINKO, Ph.D.* HANS-OLOF LISPER, Ph.D. BO EKEHAMMAR, Ph.D.

Psychology students received a 14-, 28-, or 42-hour training course in reflective listening. Before and after training, the students participated in role-played counseling conversations with confederates, who rated them. The conversations were captured on audio- or videotape, categorized, and rated by external evaluators. Results suggested that the students used reflective listening equally after different lengths of training. However, longer training resulted in the confederates disclosing more emotion, the psychology students remembering the information relayed better, and the evaluators perceiving the therapeutic relationship as better. This was especially true among the evaluators who self-reported high social skills.

REFLECTIVE LISTENING IN COUNSELING: EFFECTS OF TRAINING TIME AND EVALUATOR SOCIAL SKILLS

Many factors, such as verbal responses, personal characteristics and intentions, and context, interact to influence process and outcome in therapy and counseling (Hill, 1992). These variables are, however, difficult to investigate, and previous research has yielded few suggestions for practitioners about which type of verbal style is helpful for a certain client or problem. Moreover, there is an ongoing issue on whether client change in counseling is due to specific verbal or technical interventions, or to factors common in all psychotherapies, an issue that is essential to the training of professional helpers.

A review of studies, mainly in the cognitive-behavioral realm (Ahn & Wampold, 2001), concluded that no specific ingredients are responsible for positive outcomes in psychotherapy (the so-called Dodo bird effect),

AMERICAN JOURNAL OF PSYCHOTHERAPY, Vol. 61, No. 2, 2007

Uppsala University, Sweden. **Mailing address:* Department of Psychology, Uppsala University, Box 1225, SE-75142 Uppsala, Sweden. e-mail: Erik.Rautalinko@psyk.uu.se

and that training should focus on common factors, such as interviewing skills and the working alliance. Some things not elaborated on in the review are the specific effects *within* as well as *between* treatments. Effects on outcome between treatments are measured as differences among theoretically distinct and sometimes manualized approaches, whereas effects within treatments originate from how a treatment is given in its own theoretical framework, by a therapist who adheres to a treatment plan or a manual. It is common to find variability in outcomes among therapists, see, for example, a recent study of alliance-fostering therapy for depression (Crits-Cristophe et al., 2006). In fact, it is the effects within treatments that Ahn and Wampold (2001) studied when they concluded that variance in outcome follows variances in verbal techniques, whether the specific intervention is, for example, clarification or exposure *in vivo*.

Comprehensive descriptions of verbal responses in counseling and psychotherapy are found elsewhere (e.g., Elliott et al., 1987; Hill, 1992), but a distinction is often made between directive and nondirective responses. The latter are given special emphasis in client-centered therapy (later, person-centered therapy; Rogers, 1951, 1961) or nondirective counseling. Originally created as a reaction to directive psychoanalytic approaches, nondirective counseling became prominent in the 1940s. Nondirective counselors believe in reflecting and clarifying the messages of their clients and in the human capacity to make changes without expert advice (Phillips, 1999). In fact, Rogers claimed that simply listening to a client, especially to her or his emotions, was helpful. He clarified later (1975) that his approach had been misinterpreted as a technique, but he still maintained that empathy was not an intangible quality of a relationship, but something observeable in behavior. Rogers also promoted specific responses among those found in research on counseling, namely listening, exploring, and supporting. We suggest that these responses may explain some of the common effects in therapy and counseling.

Reflective listening (RL), a cornerstone of Rogerian nondirective counseling, means that a therapist grasps "what it is the sender is feeling or what his message means. Then he puts his understanding into his own words (code) and feeds it back for the sender's verification" (Gordon, 1970, p. 50). This description focuses on overt behavior instead of intention, and it has met resistance from investigators who believe that an intangible working alliance is therapeutic (e.g., Glauser & Bozarth, 2001). However, we welcome descriptions that simplify the process of counseling. Although not necessarily theoretically valid descriptions of a complicated process, they can make the education of counselors easier.

Reflective listening is common among many professional helpers, and it Reflective listening is common among many professional helpers, and it emphasizes that the content of another person's story should not to be altered. As Gordon (1970) put it, "The receiver does not send a message of his own—such as an evaluation, opinion, advice, logic" (p. 53). By giving such advisement, Gordon ruled out many of the verbal responses typically found in sessions conducted by counselors who, in fact, are regarded as professional by their peers (e.g., Hill et al., 1988; Nagel, Hoffman, & Hill, 1995). Gordon provided some written examples of RL, but more detailed descriptions have been given elsewhere. For example, Rautalinko and Lisper (2004) operationalized RL as seven categories of utterances:

- minimal encouraging (see Davis, 1986; Mansfield, 1991; Ralph & Thorne, 1993),
- direct encouraging (to explicitly encourage by saying "go on" or "tell me more"),
- reflecting fact.
- reflecting emotion,
- recapitulating (see Hill, 1992; Horne, Vatmanidis, & Careri, 1994; Knippen & Green, 1994; Lin, Kelly, & Nelson, 1996; Phillips, 1999), and

and • questioning on fact (open- and closed-ended) and • questioning on emotion (open- and closed-ended). Of these seven categories of verbal responses: reflecting fact, reflecting emotion, questioning on fact, and questioning on emotion comprise the core skills of RL, because they involve exploring and understanding the sender's message, and according to Gordon's (1970) definition, putting the story into the therapist's own words. Moreover, open-ended questions are more in accordance with RL because they do not restrict story telling. Hayes, Follette, and Follett (1995, pp. 159–160) remarked that a paragon of pondirective skill may be undefinable

Hayes, Follette, and Follett (1995, pp. 159–160) remarked that a paragon of nondirective skill, may be undefinable. Carl Rogers articulated that RL changes a therapeutic relationship to the positive for both interacting parties. The concept of alliance was earlier associated with psychodynamic therapies (Horvath & Luborsky, 1993), but since the introduction of self-report instruments, such as the Working Alliance Inventory (WAI, Horvath & Greenberg, 1989), the therapeutic relationship has been found to predict recovery from many psychological (Raue, Goldfried, & Barkham, 1997) and physiological problems (e.g., Horne, Vatmanidis, & Careri, 1994). Recently, the use of nondirective approaches in other psychological interventions, for example, in cognitive behavior therapy (CBT) has increased. Nondirective style has been applied

in motivational interviewing (Miller & Rollnick, 2002), dialectic behavior therapy (Linehan, 1993), and mindfulness (Kabat-Zinn, 2003), making the distinction between nondirective and directive verbal style a pantheoretical issue. Motivational interviewing (MI), for example, is described as a "client-centered, directive method for enhancing intrinsic motivation to change by exploring and resolving ambivalence" (retrieved from http:// www.motivationalinterview.org, on December 4th, 2006). This means that MI blends the RL techniques of showing empathy, with therapist intention, for example, when a therapist believes a client should become drug free.

Carl Rogers stated that a therapist's ability to be empathic and unconditionally accepting is a necessary and sufficient condition for therapeutic gain. Other more technique-focused investigators (e.g., Nugent & Halvorson, 1995) have thus drawn parallels between Rogerian empathic listening and sub-skills of RL. Yet, not much is known about the techniques that improve the therapeutic relationship (Horvath & Luborsky, 1993) or the elements that build working alliance (Horvath, 2006). Therefore, studies investigating specific interventions, and how they contribute to the alliance, are needed. The purpose of the present study was to assess the effects of training-time length on psychology students' RL counseling skills. A recent study (Rautalinko & Lisper, 2004) supports the effects of a shortterm training of RL. The present study investigated different lengths of training (of which the shortest was comparable to the training in the previous study) and the effect on RL skills among students. The major hypothesis was that RL training affects the behavior of both the psychology students and their counterparts, those who they counselled. More specifically, we expected that students who had more RL training would

- (1) use more RL responses,
- (2) receive more information from their counterparts, and
- (3) remember the information given to them better.

We also expected that students who had longer RL training time would

- (4) reach a better working alliance with their counterparts, and
- (5) receive more positive evaluations of their counseling skills.

Investigating the effects of different training times is important because skills training is popular but also costly. When time or money is limited, it may be necessary to weigh costs against expected outcomes.

TRAINING IN REFLECTIVE LISTENING

This first part of the study investigated hypotheses (1), (2), (3), and (4). In focus were psychology students, one of many populations that train

	14-hour		28-hour		42-hour		Time
RL core responses	Pretest	Posttest	Pretest	Posttest	Pretest	Posttest	F(1, 20)
Reflecting fact	3.3 (2.7)	5.2 (0.9)	3.2 (1.3)	6.4 (2.1)	3.4 (2.0)	7.3 (2.4)	48.47***
Reflecting							
emotion	0.8 (1.1)	2.2 (1.0)	1.6 (0.7)	3.0 (1.4)	1.1 (1.0)	4.7 (2.5)	25.04***
Question on fact	8.5 (1.4)	5.7 (2.5)	8.4 (3.1)	2.8 (2.2)	9.1 (2.7)	3.3 (2.6)	34.84***
Question on							
emotion	2.3 (1.5)	1.6 (1.4)	2.3 (1.1)	0.7 (0.9)	2.8 (1.6)	0.3 (0.4)	31.89***

Table I.MEAN SCORES (SDS IN PARENTHESES) OF STUDENTS' RL CORE
RESPONSES AT PRE- AND POSTTEST, AND MAIN EFFECTS OF TIME
IN A REPEATED MEASURES ANOVA

***p < .001.

counseling skills. The first and second authors were, at the time, responsible for training at university level.

Method

Participants

The participants were undergraduate psychology students who enrolled in the study during their third semester of study. There were 7 male and 21 female participants, with a mean age of 25 years (range = 21-41years), and they were randomly assigned to one of the three RL training conditions. At the onset of the study, there were another six participants, who, for various reasons, could not fulfil the study. When the study was completed, attendance for the three RL training conditions were: 14 hours (n = 9, attrition = 2), 28 hours (n = 12, no attrition), and 42 hours (n = 12, no attrition)7, attrition = 4). As attrition was greatest in the most demanding training group, we investigated if the participants who dropped out were the least skilled or motivated from the start of training. We analysed attrition, comparing the participants who dropped out and the remaining participants by age, pretraining dependent variables (see Table 1), and scores on the Working Alliance Inventory (described below) using unrelated t tests. No statistically significant differences were found (ts varying between 0.11 and 1.04, ps varying between .23 and .79), and the direction of the differences between the drop-out and remaining participants was not systematic. Thus, the drop-out participants were excluded from subsequent analyses.

A 26-year-old female and a 28-year-old male confederate were paid to act as clients in evaluations before and after training. One of the confederates was a university student, but not a psychology major, the other person had neither a university education nor affiliation. None of the confederates had had any counseling skills training.

Procedure

Training. The students were randomly assigned to one of three training conditions. The 14-hour training consisted of modelling and role-play in groups of 3 to 5 participants during seven 2-hour sessions over a time-span of two weeks. The training aimed at teaching the participants to use nondirective responses such as RL, trainers were the first and second author. The students in the 28-hour and 42-hour training conditions had the same module of training as the students in the 14-hour condition, but repeated once (28-hour condition) or twice (42-hour condition). The intermission between the trainings was approximately a month, and there was no other counseling skills training during the time.

The training times (factors of a 14-hour module) were chosen for two reasons. First, the shortest one is comparable to a training used in a recent study (Rautalinko & Lisper, 2004), considered short but effective. Second, the medium-long training is comparable to the time length of the most common nondirective counseling skills training, namely 24 hours (Rinn & Markle, 1977).

Conversations with confederates. A week before the first training session and a week after the last, all students had 7-minute, audiotaped, role-played counseling conversations with the confederates acting as clients. The students were instructed to treat the confederates as they would treat a client, and the confederates were blind to the students' group affiliations. The four roles (clinical case descriptions; American Psychiatric Association, 1994) were balanced over pre- and posttest so that the confederates acted all roles and no student met with a confederate acting a specific role more than once. The four roles were related to the following problem domains: (1) post-traumatic stress, (2) depression, (3) hyperactivity, and (4) binging. The confederates covered some complaints encountered by mental health and personal counselors in the meetings, but the roles did not include specific, scripted responses by the confederates.

Confederate ratings. After each conversation, the confederates filled out the Working Alliance Inventory (WAI; Horvath & Greenberg, 1989), which consists of 36 statements answered on a 7-point scale ranging from 1 (*never*) to 7 (*always*). The WAI includes three subscales that assess bonds, goals, and tasks, according to a three-partite model by Bordin (1979; see also Hatcher & Gillaspy, 2006). The WAI gives a composite score between 36 and 252, where a high score means that a counselor and

a client agree on the problem areas to be worked on, and implies trust and approval in the relationship.

Student ratings. The students filled in a self-report instrument that measured recall of what the confederate had said during the counseling. It consisted of a 15-item checklist of true–false statements that described the role the confederate had acted. A high score on this recall of information checklist meant that the student remembered information given to her or him by the confederate.

Response categorization. The conversations were transcribed, and the middle 5-minute part of each conversation was categorized into RL and other responses, and the conversations were quantified. The first and last minute of all conversations, including greetings and goodbye-responses, which would have been difficult to categorize according to the hypotheses, were excluded. An equal length of each conversation was transcribed. Verbal behavior coding was done by content analysis (Holsti, 1969; Krippendorff, 1980; Weber, 1990), which meant developing categories that were mutually exclusive and exhaustive. We categorized conversations by coding units, defined as one member of the dyad speaking before the other responded. We coded repeated content in a unit only once. The categories were either anticipated from theory or previous research, or new categories emerged during the analysis. The RL category choices determined before the conversation analyse and found in the students' responses were: minimal encouraging, direct encouraging, reflecting fact, reflecting emotion, questioning on fact, and questioning on emotion. The aforementioned category recapitulation requires longer conversations than the present role-plays could provide. The non-RL categories that emerged were advice and interpretation. The categories corresponded to those used in another study of psychology students' counseling skills (Lisper & Rautalinko, 1996). The categories that emerged for the confederates' responses were minimal encouraging, disclosing fact, disclosing emotion, answering question on fact, answering question on emotion, posing question on fact, and posing question on emotion.

The first author coded all transcripts, and interrater reliability was calculated for four randomly chosen transcripts over all three conditions at pre- and posttest, which were coded independently by the first and second authors. The Cohen Kappa coefficients ranged between .82 and 1.00 for the students' responses (n = 8), and between .73 and .93 for the confederates' responses (n = 7), which was considered acceptable (Hill, 1991). Category drift for the first author's coding was calculated for four other randomly chosen transcripts, which were coded anew by the first author

after 16 months. The Kappa coefficients for these two codings ranged between .89 and 1.00 for the students' responses, and between .87 and 1.00 for the confederates' responses.

Design and analysis. The design was experimental, with participants randomly assigned to one of the three training conditions. A 3×2 (Condition: 14-, 28-, and 42-hours of training RL; time: pre- and posttest) repeated measures MANOVA was used to test the effects of training time length on the participants' RL core responses. In case of statistical significance, the included four response categories were further investigated in 3×2 repeated measures ANOVAs. The differences in the students' responses could have been tested in $3 \times 2 \times 2$ (Condition \times Time \times Confederate) ANOVAs as well, but possible interaction effects of confederates would be difficult to interpret as the confederates may differ on any confounding variable. That is why both confederates acted all four roles and the students' and the two confederates' responses were averaged for all 5-minute sequences of audiotaped conversations.

RESULTS AND COMMENTS

Students' Responses

Table 1 presents the students' use of RL core responses at pre- and posttest. Prior to training, the students posed many questions on fact (overall M = 8.6). Some of the factual content given by the confederates was paraphrased, shown by the category reflecting fact (overall M = 3.3). After training, however, reflecting fact was used more often (overall M = 6.2) than questioning on fact (overall M = 3.9).

A repeated-measures MANOVA of the RL core responses resulted in a statistically significant main effect of time between pre- and posttest (Wilk's $\Lambda = .22$, F(2, 24) = 43.62, p < .001), but a nonsignificant Time × Condition interaction effect (Wilk's $\Lambda = .92$, F(4, 48) = 0.50, p = .74). Results of the separate-repeated measures ANOVAs of the RL core responses showed that the use of each had changed between pre- and posttest. The nonsignificant Time × Condition interaction effect did not support that the effects of training would be larger for the students who train for a longer time.

In conclusion, although the students in all training groups significantly increased their use of RL core responses, the results of the analyses do not support hypothesis (1).

Confederates' Responses

The two confederates' responses were averaged for the 5-minute parts of conversations as well. Both at pre- and posttest, the confederates

disclosed about twice as much fact (overall M = 24.3) than emotion (overall M = 11.5). They answered more questions on fact (overall M = 5.4) than on emotion (overall M = 1.0), a result that corresponds to the students' posing more questions on fact than on emotion.

The confederates used somewhat different responses at pre- and posttest. There was a significant main effect of time (between pre- and posttest) on disclosing emotion, as the confederates disclosed more emotion at posttest (overall M = 12.9) than at pretest (overall M = 10.0), F(1, 20) =6.60, p < .05. None of the Condition \times Time interaction effects on the confederate responses (disclosing fact or emotion; posing question on fact or on emotion) was statistically significant (Fs varying between 0.14 and 2.42, ps between .11 and .87). Our interpretation of these results was that the confederates did not speak differently to students with various amounts of training.

In conclusion, the results of the analyses do not support Hypothesis (2).

Students' Recall of Information

The students' recall of information given by the two confederates was averaged for 5-minute conversations both pre- and posttest. Only questions where information was found in the transcripts were included in the analysis. Two out of three training groups had higher recall scores after training (overall M = 9.7) than before (overall M = 9.5). However, this main effect of time was not statistically significant in a repeated measures ANOVA, F(1, 19) = 1.19, p = .29. Yet, there was a significant Condition \times Time interaction effect, F(2, 19) = 5.93, p < .01, in favor of the 42-hour training group, which had the highest increase in their mean recall scores. This mean increase was 1.20 units (overall SD = 0.65 before and after training), compared to 0.55 (SD = 1.25) units for the 28-hour condition, and -0.93 units (SD = 0.65) for the 14-hour condition.

In conclusion, the results of the analyses support hypothesis (3) student with more training remember the information given to them better.

Confederates' Ratings

The Cronbach coefficient alpha of the WAI was .88, in line with previously reported reliability data (Horvath & Greenberg, 1989). The WAI scores given by both confederates to a student were averaged. The confederates' WAI scores were lower post-training for all training groups (overall M = 171.7 after, overall M = 182.8 before), but the difference was statistically nonsignificant in a repeated measures ANOVA, F(1, 20) = 0.43, p = .52. The Condition × Time interaction effect was nonsignificant as well, at F(2, 20) = 2.13, p = .14.

In conclusion, the results of the analyses do not support hypothesis (4).

EXTERNAL EVALUATION

This second part of the study investigated hypothesis (5), which states that students who have trained RL for a longer time receive more positive evaluations of their counseling skills than students who have trained RL for a shorter time. This evaluation was based on the results of a larger sample of raters instead of the two confederates. Moreover, we examined if there would be a relationship between the raters' evaluations and their own level of social skills. Professionals working with intellectually or socially challenged clients sometimes find that directive responses, such as instructions combined with praise, are more appropriate than letting clients decide for themselves what steps to take. There is also previous evidence that client functioning can be related to the verbal responses of a counsellor (e.g., Hill et al., 1988), and we hypothesized that raters with good social skills would appreciate a nondirective style.

Method

Participants

The 43 participants in an external validation sample consisted of 26 vocational counselors (10 male and 16 female, mean age 43.2 years, range = 31-64 years; mean professional experience 18.7 years, range = 2-41 years) and 17 psychology students (6 male and 11 female, mean age 29.6 years, range = 22-44 years) who had had no counseling skills training. The authors acknowledge that the sample consisted of participants from merely two populations, but at least they had different degrees of experience with respect to counseling.

Procedure

The evaluators watched and rated 18 videotaped counseling conversations, six randomly selected conversations from each of the three training groups (14-, 28-, and 42-hours). These conversations were videotaped 12 months after the training sessions, and there had been no additional RL training of the students. Instead, the students were instructed to videotape a 10 to 20 minute role-played counseling conversation with senior student confederates (mean age 32 years, range = 25–52 years). The confederates' roles were related to the following problem domains: (1) spider phobia, (2) marital discord, (3) domestic violence, and (4) insomnia, and the roles did not specify any verbal responses. The mean length of the videotaped counseling conversations was 15.4 minutes (range = 10.7-22.5 min).

The external evaluators who were blind to the students' group affilia-

tion, evaluated the 5- to 10-minute interval in each conversation. The videotaped role plays were presented in a balanced order to avoid systematic sequential variance between the conditions. The external evaluators rated the videotaped students for the variables of competence, confidence, and compassion using a 10-point scale (1 = not at all to 10 = very much). The aforementioned variables describe positive characteristics of a counsellor as used in a recent study by Rautalinko and Lisper (2004). In the present study, the evaluators gave their ratings five times during each conversation, on signals given at one minute intervals. The final evaluation score was the mean score of these five ratings, ranging from 1 to 10. An average score was chosen to avoid primacy or recency effects that might affect the evaluations. These average scores on competence, confidence, and compassion given by the raters to each student "counselor" were highly related (r = .59 or above) in the three learning session conditions $(p_{\rm s} < .05, n = 43)$, and averaged to a positive evaluation index ranging from 1 to 10. The Cronbach coefficient alpha of the positive evaluation index was .92.

An open-ended measure was included as well. After having watched each videotaped conversation, the evaluators wrote down their opinion of the psychology student ("I think the counselor was. . ."). Evaluators spent approximately one minute on this task between each conversation. The authors considered open-ended measures important because we believed we could thus explore unexpected effects of longer RL skills training.

After having watched and evaluated the videos, the evaluators filled out the Social Skills Inventory (SSI; Riggio, 1986). It is a 90-item assessment of six basic social skill dimensions: sending, receiving, and controlling communication in two separate domains, an emotional-nonverbal domain and a social-verbal domain. The items were answered on a 5-point Likert scale ranging from 1 (*not at all like me*) to 5 (*exactly like me*), and the total score (theoretical range: 90 to 450), approximates the respondent's global social skills, with a higher score meaning more socially skilled.

Results and Comments

The raters' SSI scores varied between 248 and 335 (M = 292, SD = 23.8), and by a median split, the raters were categorized into either high or low in social skills. The Cronbach coefficients alpha of the SSI subscales ranged between .64 and .86. Alpha for the SSI composite score was .84, and considered satisfactory.

Positive Evaluation Index. Figure 1 presents the raters' evaluation of the students in the three training conditions. The mean positive evaluation

Figure 1

MEAN SCORES ON THE POSITIVE EVALUATION INDEX GIVEN TO THE STU-DENTS IN THE THREE TRAINING CONDITIONS BY EVALUATORS WITH HIGH AND LOW SOCIAL SKILLS



index scores given by the high vs. low social skills raters were: 5.9 vs. 5.9 (to 14-hour), 6.1 vs. 5.5 (to 28-hour), and 6.5 vs. 6.1 (to 42-hour). A two-way mixed ANOVA showed a statistically significant main effect of time, F(2, 41) = 10.86, p < .001. A Tukey HSD post-hoc test on this main effect showed that the students who had trained RL for 42 hours received significantly (p < .01) higher scores on the positive evaluation index than the students in both the 14- and the 28-hour training groups. Furthermore, the ANOVA showed a tendency to a Time × Social skills interaction effect, F(2, 41) = 3.11, p = .05. The character of this effect is somewhat difficult to describe but can be observed in Figure 1. It shows that there was no clear trend in the evaluations given by the evaluators who were low on social skills, whereas evaluators who were high on social skills gave higher evaluations to the students who were part of the group who had been trained for a longer period.

A median-split of the SSI scores can be questioned. It is arbitrary to label participants as high and low on social skills based on scores in a sample, and this leads to a loss of power as a continuous variable is dichotomized. Thus, to test the robustness of the results above, we also included the SSI as a continuous predictor variable in a two-way mixed ANOVA (Statistica software: General Linear Model). The conclusions drawn from that analysis, and the one presented above, are the same, but the latter visualizes the character of the interaction more clearly. In conclusion, the results support hypothesis (5), that students with more training would receive more positive evaluations of their counseling skills.

Open-ended descriptions

A total of 2,775 open-ended descriptions of the 18 psychology students were categorized by the first author.A 38-year-old female philologist with no more than common knowledge of psychology, who was blind to the purpose of the study, categorized a randomly selected subset of 216 descriptions. The descriptions were categorized as positive (+; for example, "helpful"), negative (-; for example, "in a hurry"), or neutral (for example, "careful"). Cohen's Kappa (n = 3) for the first author's and the philologist's categorizations was .79, which was considered acceptable. The first author's further content analysis of the 1,334 positive valence descriptions showed that approximately 35% belonged to a subcategory "interested-involved-listening", 15% to "calm-patient", and 15% to "empathic-understanding". Of the 1,193 negative valence descriptions, approximately 30% belonged to a subcategory "non-encouraging-vagueexpressionless", 30% to "insecure-tense", and 10% to "leading-managing". The remaining 248 neutral valence descriptions were not analyzed in detail, but most of them concerned body-posture or facial features. Descriptions of non-verbal aspects were not considered germane to the present study.

A further quantification was made of the positive- and negative-valence descriptions. The mean number of total (+) and (-) descriptions given by the evaluators high vs. low in social skills to each training condition is illustrated in Figure 2. The evaluators who were rated high in social skills gave a mean of 1.68 (+) and 1.77 (-) to the 14-hour, 1.98 (+) and 1.82 (-) to the 28-hour, and 2.26 (+) and 1.82 (-) to the 42-hour students. The evaluators who were rate low in social skills gave a mean of 1.61 (+) and 1.89 (-) to 14-hour, 1.50 (+) and 1.77 (-) to 28-hour, and 1.83 (+) and 1.53 (-) to 42-hour students. A two-way mixed ANOVA showed a statistically significant main effect of time for the positive descriptions, F(2,(41) = 11.93, p < .001. A Tukey HSD post-hoc test on this main effect showed that the students who had trained in RL for 42 hours received significantly (p < .01) more positive descriptions than the students in both the 14- and the 28-hour training groups. Moreover, the ANOVA showed a statistically significant Time × Social skills interaction effect for the positive descriptions, F(2, 41) = 3.38, p < .05. Figure 2 shows that evaluators high in social skills gave more positive descriptions to students

Figure 2

MEAN NUMBER OF POSITIVE DESCRIPTIONS GIVEN TO THE STUDENTS IN THE THREE TRAINING CONDITIONS BY EVALUATORS WITH HIGH AND LOW SO-CIAL SKILLS



with more RL training whereas the other evaluators' number of positive descriptions showed no clear trend. This result is very similar to that obtained for the positive evaluation index.

The two-way mixed ANOVA on the number of negative descriptions showed no statistically significant differences between the training conditions, or between evaluators with high and low social skills.

In conclusion, the results support hypothesis (5), that students with more training would receive more positive evaluations of their counseling skills

DISCUSSION

The results show that the students in all training groups increased their reflective listening skills. It is not possible to conclude whether the training led to changes in how the confederates talked about their problems, as the experiments did not include a no-training control group. This is the study's main weakness. However, it appears as though short training of communication skills may bring about the same behavioral changes as longer one, meaning that sometimes it is possible to learn a well-defined skill, such as RL, with short training. Yet, as shown by the results of the first part of the study, there were some changes in the students who had trained in RL for a longer time. The ratings were mostly in favor of a longer training time both in the working alliance as perceived by the two confederates and the evaluations made by a larger external sample. We observed this effect for both the positive evaluation index and the open-ended descriptions. As these changes were not manifested in the number of RL core responses in transcripts, perhaps they existed as prosodic differences indicating that the students who had trained longer were more comfortable with their new skills. It has been proposed that counseling skills develop through four stages: from unaware incompetence, through aware incompetence and effortful competence, to automatic competence (Lindh & Lisper, 1990). Perhaps the external evaluators in the present study sensed a prosodic difference between the effortful and the automatic stages. No difference existed in transcripts-the students knew their RL regardless of training group— but differences were perhaps discernible to an audience, especially one high on social skills. This is in line with the proposition that adopting a reflective listening style (not to mention a nondirective counseling format) is not an easy task, and is perhaps much more complicated than a transcript of a counseling session could indicate (Wickman & Campbell, 2003).

The results of the last part of the study showed an interaction between length of training time and the evaluators' social skills. Participants who reported above average social skills for the sample evaluated the students in the 42-hour training condition more positively than the students who had less training. This effect must be interpreted with caution, as there was a restriction of range in social skills in the present sample. The average SSI scores for the low and high social skills groups were 271 and 311 respectively, which gives a difference of 40 points. Riggio (1989) reported higher cutoff scores for low and high social skills participants, and also a larger difference between them. The result is, nevertheless, theoretically valuable, and it may have practical significance as well. We do not advocate using a measure of social skills to allocate clients into different types of counseling, as has been done earlier (e.g., Helweg & Gaines, 1977). But the results of this study suggest that when working with socially skilled clients, the counsellor should acknowledge that her or his verbal style is under increased scrutiny. Why this is the case may be investigated in future research. Hypothetically, the socially skilled evaluators may identify with the counsellors' (trained) skills as being similar to their own, and identification may increase liking. But then training would raise the psychology students' social skills, which seems less likely. Another possibility is that the socially skilled evaluators are also more cognitively capable, being able to attend to both parties in the conversation simultaneously. If an evaluation score of a counseling session depends on the behaviour of both the client and the counsellor, then it is possible that a more cognitively capable

evaluator has a better chance of recognizing differences between counsellor groups. A third possibility is that the socially skilled evaluators simply recognize subtle verbal behaviours that they know are considered socially desirable. Awareness of social codes is one aspect of social competence.

The results of the recall test showed that students who had trained RL for a longer time remembered information given to them better than students who had trained for a shorter time. This suggests that more training keeps the listener on track, even if the design in the present study cannot answer whether this is an effect of enhanced RL skills or an unspecific effect of more role playing. A study of repetition and selfassessment of questions (Parkin, Wood, & Aldrich, 1988) showed that recall of orally presented information is influenced by a time lag, recall is better when other information. This may imply that a reflective listener recalls more information if she or he paraphrases the other person's story a while after hearing. Investigating effects of time lag in RL is a future research possibility.

The design of the first part of the study would have benefited from replacing one of the three present conditions with a no-training control group. Previous research, however, has shown beneficial effects of even very short training in counseling skills (i.e., 11 sessions in Ali et al., 2003; 16 hours in Rautalinko & Lisper, 2004). Another methodological limitation of the present study is the small number of participants. Practical circumstances limited the number of participants in each condition: inclusion and attrition could not be controlled, and the authors chose to omit all data on participants who did not fulfil their participation in the study. We wish to acknowledge that a small n diminishes statistical conclusion validity, and makes the results of the study more difficult to generalize. However, we had no indication that the present sample of psychology students would differ from groups of students who are trained each year using the same method.

The skills used in nondirective counseling date back at least to the 1940s. Since then, there has been a growing body of research on the methods described by Carl Rogers and others (e.g., Gordon, 1970), and to related concepts such as empathy. The earliest criticism of nondirective counseling concerned the question whether it should be kept separate from psychotherapy. Hathaway (1948/2000), for example, proposed that rapport is a common feature to all types of counseling or psychotherapy, and not an aspect unique to nondirective counseling. He insisted that nondirective counseling be evaluated on progress-oriented grounds, in-

stead of focusing on process in interviews or by demonstrating positive correlates to treatment. His remark "... almost any form of attention given to a patient's problems, whether skilled and upon good theoretical basis or unskilled and poorly conceived, is likely to result with improvement" (p. 858) describes the threat that the interest in nondirective counseling posed to adherents to directive psychotherapy. For those who value rigorous theory about the inner life of clients, the fact that seemingly easy verbal exhortations could produce personal growth seems like a hoax. But Hathaway's remark describes something else as well. More recent investigations demonstrated that warmth and support often explain a considerable part of the variance in client recovery, for example from depression, but also that a considerable part of recovery depends on therapeutic techniques (Burns & Auerbach, 1996). It may be that the therapist's emphasis on one technique is determined by her or his theoretical orientation, although some investigators (e.g., Kelly, 1997) proposed that there is an interdependence between relationship/humanistic and task/technical dimensions in all counseling, regardless of theoretical orientation.

Those interested in counseling research have witnessed disagreement on which techniques should be used (for example, experiential methods vs. transference clarification), and on whether a good therapeutic relationship is a sufficient—or even necessary—condition for client recovery. There is no need to widen the gap between proponents of relationshiporiented and goal-oriented treatments. Neither do we propose that short training automatically produces counselors. Instead, we suggest that a combination of good_role models, training, self-reflection, and experience—of which the present study has explored short training of counseling skills—are all needed in the education of professional helpers.

REFERENCES

- Ahn, H., & Wampold, B.E. (2001). Where oh where are the specific ingredients? A meta-analysis of component studies in counseling and psychotherapy. *Journal of Counseling Psychology*, 48, 251-257.
- Ali, B.S., Rahbar, M.H., Naeem, S., Gul, A., Mubeen, S., & Iqbal, A. (2003). The effectiveness of counseling on anxiety and depression by minimally trained counselors: A randomized controlled trial. *American Journal of Psychotherapy*, 57, 324–336.
- American Psychiatric Association (1994). DSM-IV casebook. Washington, DC: Author.
- Bordin, E.S. (1979). The generalizability of the psychoanalytic concept of the working alliance. Psychotherapy: Theory, Research, and Practice, 16, 252-260.
- Burns, D.D., & Auerbach, A. (1996). Therapeutic empathy in cognitive-behavioral therapy: Does it really make a difference? In P.M. Salkovskis (Ed.), *Frontiers of cognitive therapy* (pp. 135–164). New York: Guilford Press.
- Cedar, B., & Levant, R.F. (1990). A meta-analysis of the effects of Parent Effectiveness Training. American Journal of Family Therapy, 18, 373-384.
- Crits-Christoph, P., Connolly-Gibbons, M.-B., Crits-Cristophe, K., Narducci, J., Schamberger, M., &

Gallop, R. (2006). Can therapists be trained to improve their alliances? A preliminary study of alliance-fostering psychotherapy. *Psychotherapy Research*, *16*, 268–281.

- Davis, K. (1986). The process of problem (re)formulation in psychotherapy. Sociology of Health and Illness, 8, 44-74.
- Elliott, R., Hill, C.E., Stiles, W.B., Friedlander, M.L., Mahrer, A.R., & Margison, F.R. (1987). Primary therapist response modes: Comparison of six rating systems. *Journal of Consulting and Clinical Psychology*, 55, 218–223.
- Glauser, A.S., & Bozarth, J.D. (2001). Person-centered counseling: The culture within. Journal of Counseling and Development, 79, 142–147.
- Gordon, T. (1970). PET: Parent Effectiveness Training. New York: Wyden.
- Gordon, T. (1977). LET: Leader Effectiveness Training. New York: Wyden.
- Hatcher, R.L., & Gillaspy, J.A. (2006). Development and validation of a revised short version of the Working Alliance Inventory. *Psychotherapy Research*, 16, 12–25.
- Hathaway, S.R. (1948/2000). Some considerations relative to nondirective counseling as therapy. Journal of Clinical Psychology, 56, 853–859.
- Hayes, S.C., Follette, W.C., & Follette, V.M. (1995). Behavior therapy: A contextual approach. In A.S. Gurman & S.B. Messer (Eds.), *Essential psychotherapies: Theory and practice* (pp. 128–181). New York: Guilford Press.
- Helweg, G.C., & Gaines, L.S. (1977). Subject characteristics and preferences for different approaches to psychotherapy: A multivariate study. *Journal of Consulting and Clinical Psychology*, 45, 963–964.
- Hill, C.E. (1991). Almost everything you ever wanted to know about how to do process research on counseling and psychotherapy bud didn't know who to ask. In C. Watkins & L. Schneider (Eds.), *Research in counseling* (pp. 85–118). Hillsdale, NJ: Erlbaum.
- Hill, C.E. (1992). Research on therapist techniques in brief individual therapy: Implications for practitioners. *Counseling Psychologist*, 4, 689-711.
- Hill, C.E., Helms, J.E., Tichenor, V., Spiegel, S.B., O'Grady, K.E., & Perry, E.S. (1988). Effects of therapist response modes in brief psychotherapy. *Journal of Counseling Psychology*, 35, 222– 233.
- Holsti, O.R. (1969). Content analysis for the social sciences and humanities. Reading, MA: Addison-Wesley.
- Horne, D.J. de L., Vatmanidis, P., & Careri, A. (1994). Preparing patients for invasive medical and surgical procedures. 2: Using psychological interventions with adults and children. *Behavioral Medicine*, 20, 15–21.
- Horvath, A.O. (2006). The alliance in context: Accomplishments, challenges, and future directions. *Psychotherapy: Theory, Research, Practice, Training, 43, 258–263.*
- Horvath, A.O., & Greenberg, L.S. (1989). Development and validation of the Working Alliance Inventory. *Journal of Counseling Psychology*, 36, 223-233.
- Horvath, A.O., & Luborsky, L. (1993). The role of the therapeutic alliance in psychotherapy. Journal of Consulting and Clinical Psychology, 61, 561-573.
- Kabat-Zinn, J. (2003). Mindfulness-based interventions in context: Past, present, and future. Clinical Psychology: Science and Practice, 10, 144–156.
- Kelly, E.W. Jr. (1997). Relationship-centered counseling: A humanistic model of integration. Journal of Counseling and Development, 75, 337-345.
- Knippen, J.T., & Green, T.B. (1994). How the manager can use active listening. Public Personnel Management, 23, 357-359.
- Krippendorff, K. (1980). Content analysis: An introduction to its methodology. Beverly Hills, CA: Sage.
- Lin, M.-J., Kelly, K.R., & Nelson, R.C. (1996). A comparative analysis of the interpersonal process in school-based counseling and consultation. *Journal of Counseling Psychology*, 43, 389–393.
- Lindh, G., & Lisper, H.-O. (1990). Samtal för förändring [Conversation that changes]. Lund, Sweden: Studentlitteratur.
- Linehan, M.M. (1993). Cognitive-behavioral treatment of borderline personality disorder. New York: Guilford.
- Lisper, H.-O., & Rautalinko, E. (1996). Effects of a six hour training of active listening. Scandinavian Journal of Behaviour Therapy, 25, 117–125.
- Mansfield, F. (1991). Supervised role-play in the teaching of the process of consultation. Medical

Education, 25, 485-490.

- Miller, W. R., & Rollnick, S. (Eds.) (2002). *Motivational interviewing: Preparing people for change* (2nd ed.). London: Guilford.
- Nagel, D. P., Hoffman, M. A., & Hill, C. E. (1995). A comparison of verbal response modes used by master's-level career counselors and other helpers. *Journal of Counseling and Development*, 74, 101–104.
- Nugent, W.R., & Halvorson, H. (1995). Testing the effects of active listening. Research on Social Work Practice, 5, 152–175.
- Parkin, A.J., Wood, A., & Aldrich, F.K. (1988). Repetition and active listening: The effects of spacing self-assessment questions. *British Journal of Psychology*, 79, 77–86.
- Phillips, B. (1999). Reformulating dispute narratives through active listening. Mediation Quarterly, 17, 161–180.
- Ralph, A., & Thorne, E. (1993). Defining and enhancing competent interview behaviour using the Verbal Interaction Analysis System. Scandinavian Journal of Behaviour Therapy, 22, 65–87.
- Raue, P.J., Goldfried, M.R., & Barkham, M. (1997). The therapeutic alliance in psychodynamicinterpersonal and cognitive-behavioral therapy. *Journal of Consulting and Clinical Psychology*, 65, 582–587.
- Rautalinko, E., & Lisper, H.-O. (2004). Effects of training reflective listening in a corporate setting. Journal of Business and Psychology, 18, 281–299.
- Riggio, R.E. (1986). Assessment of basic social skills. Journal of Personality and Social Psychology, 51, 649-660.
- Riggio, R.E. (1989). Social Skills Inventory Manual: Research edition. Palo Alto, CA: Consulting Psychologists Press.
- Rinn, R.C., & Markle, A. (1977). Parent effectiveness training: A review. Psychological Reports, 41, 95-109.
- Rogers, C.R. (1951). Client-centered therapy. Cambridge, MA: Riverside Press.
- Rogers, C.R. (1961). On becoming a person: A therapist's view of psychotherapy. London: Constable.
- Rogers, C.R. (1975). Empathic: An unappreciated way of being. Counseling Psychologist, 5, 2-10.
- Weber, R.P. (1990). Basic content analysis (2nd ed.). Newbury Park, CA: Sage.
- Wickman, S.A., & Campbell, C. (2003). An analysis of how Carl Rogers enacted client-centered conversation with Gloria. Journal of Counseling and Development, 81, 178–184.