

See discussions, stats, and author profiles for this publication at: <https://www.researchgate.net/publication/232507178>

Helping Skills Training for Undergraduates: Outcomes and Prediction of Outcomes

Article in *Journal of Counseling Psychology* · July 2008

DOI: 10.1037/0022-0167.55.3.359

CITATIONS

58

READS

3,151

6 authors, including:



Clara Hill

University of Maryland, College Park

382 PUBLICATIONS 17,438 CITATIONS

SEE PROFILE

Some of the authors of this publication are also working on these related projects:



Training Undergraduate Students to Use Insight Skill [View project](#)

Helping Skills Training for Undergraduates: Outcomes and Prediction of Outcomes

Clara E. Hill, Melissa Roffman, Jessica Stahl, Suzanne Friedman, Ann Hummel, and Chrisanthy Wallace
University of Maryland

The authors examined outcomes and predictors of outcomes for 85 undergraduates in 3 helping skills classes. After training, trainees used more exploration skills in helping sessions with classmates (as assessed by perceptions of helpees and helpers/trainees as well as behavioral counts of skills), were perceived by helpees as more empathic, talked less in sessions, conducted better sessions (from helpee and helper/trainee perspectives), and reported higher self-efficacy for using helping skills. In addition, trainees' confidence increased while learning exploration skills, dropped while learning insight skills, and then increased again while learning action skills. The authors were not able to predict outcome from the variables used (grade-point average, empathic concern and perspective taking, perfectionism). Suggestions for training and future research on training are included.

Keywords: helping skills training, counselor development, exploration skills, empathy

In the 1960s and 1970s, numerous training programs were developed on the basis of Rogers' (1942, 1951) client-centered therapy to train professionals to use helping skills (see reviews in Matarazzo, 1971, 1978; Russell, Crimmings, & Lent, 1984). This training quickly spread to training peer counselors (Aiken, Brownell, & Iscoe, 1974; Leventhal, Berman, McCarthy, Wasserman, 1976) to provide services on campuses and in communities. In addition, training became popular for undergraduate students (Cowen, Gardner, & Zax, 1967; Guerny, 1969), with the assumption that such training would prepare students going on for professional degrees as well as increase the marketability of those not continuing to graduate degrees (Korn, 1980). Unfortunately, not much research has been conducted on the effects of helping skills training for undergraduate students. Thus, the purpose of this study was to investigate the outcomes and the predictors of outcome of helping skills training for undergraduate psychology students, as was recommended recently by Hill and Lent (2006).

Empirical Research on Helping Skills Training

Hill and Lent (2006) reviewed both the narrative and meta-analytic reviews of the literature on helping skills training. Narrative reviews (Ford, 1979; Kasdorf & Gustafson, 1978; Matarazzo, 1971, 1978; Russell et al., 1984) concluded that warmth and empathy could be taught to both lay and professional personnel. Baker, Daniels, and Greeley (1990) in their meta-analysis reported

large (1.07), medium (.63), and small (.20) effect sizes for Carkhuff's (1969) human relations training, Ivey's (1971) micro-counseling, and Kagan's (1984) interpersonal process recall, respectively for graduate-level trainees. There may have been a confound between type of training and length of training, however, given that effect sizes were higher from programs that had more hours of training (i.e., the average length of training was 37, 19, and 9.5 hr for Carkhuff's, Ivey's, and Kagan's approaches, respectively).

We found only five studies focused on training undergraduate students, with four of the five studies finding positive effects for training. Both Payne and Woudenberg (1978) and Korn (1980) found effects for training when students' written responses to written analogue presentation of client stimuli were coded by trained judges using Carkhuff's (1969) Empathy scale. In contrast to expectations, Volz, Klevans, Norton, and Putens (1978) found that control participants (students in other human services classes) actually performed better in terms of interpersonal communication skills (as assessed through interviews with coached clients) after training than did trainees in communication skills training. In a follow-up study, Klevans, Volz, and Friedman (1981) found that students involved in an experiential program where they practiced skills used significantly more facilitative skills in a helping session with a coached client than did students who just observed and analyzed skills in clinical interactions. Klevans et al. attributed the difference in findings in this study from the earlier Volz et al. study to more time in training, which supports the meta-analytic results cited above by Hill and Lent (2006) that more training may lead to better outcomes. More recently, Hill and Kellems (2002) assessed the effects of helping skills training in undergraduate helping skills classes, using ratings by helpees in brief helping sessions. Helpees viewed their undergraduate helpers as using more exploration skills, insight skills, and action skills; fostering a better therapy relationship; and conducting better overall sessions at the end than at the beginning of the semester.

Clara E. Hill, Melissa Roffman, Jessica Stahl, Suzanne Friedman, Ann Hummel, and Chrisanthy Wallace, Department of Psychology, University of Maryland

Suzanne Friedman is now at the counseling and career services, Howard Community College, Columbia, Maryland.

We thank Mark Johnson for statistical consultation.

Correspondence concerning this article should be sent to Clara E. Hill, Department of Psychology, University of Maryland, College Park, MD 20742. E-mail: hill@psyc.umd.edu

Although the Hill and Kellems (2002) study was an advance over past research in that it involved helping situations with volunteer helpees presenting real problems, it involved a limited number of outcomes and assessed only the helpee perspective.

Methodological Issues Assessing Training Outcomes

Much of the previous research on helping skills training suffered from serious methodological limitations (see Hill & Lent, 2006). Among the most problematic limitations are the use of analogue clinical stimuli and the limited number of outcomes assessed. We review these methodological issues briefly.

The Use of Analogue Stimuli

In their review of the helping skills literature, Hill and Lent (2006) indicated that most studies investigated the outcomes of helping skills training by using analogue clinical stimuli. In other words, researchers pretested participants using some sort of written or audiotaped client stimuli with instructions to respond as helpfully as possible. They then provided training on the outcome variable (i.e., empathy or reflection of feelings) in the experimental condition. Following the experimental manipulation, trainees in both the experimental condition and control condition were post-tested on the outcome variable using similar written or audiotaped client stimuli. Using analogue clinical stimuli is problematic because what is being measured is trainees' ability to produce cued responses. Given that it is much more difficult to respond in a prescribed way in the moment to a person presenting a real problem than when one can take time to respond to a scripted analogue stimulus, it is doubtful that results from analogue stimuli would generalize to actual helping situations. Given these difficulties with analogue client stimuli, Hill and Lent (2006) suggested that the best demonstration of the effects of training would be to assess trainees' performance of the skills in unscripted helping sessions.

Limited Range of Outcomes

Another problem is that most studies assessed a limited range of training outcome variables. Hill and Lent (2006) noted that almost all the studies assessed either the ability to be empathic (as measured by trained raters using the Carkhuff, 1969, or Truax and Carkhuff, 1967, rating scales) and/or the ability to produce reflection of feelings (coded by trained judges using a variety of coding systems). Hill, Stahl, and Roffman (2007) suggested that helping skills training ideally results in a number of outcomes in addition to empathy and reflection of feelings. We next speculate about some of these outcomes.

Perhaps the most fundamental measure of the effectiveness or outcome of helping skills training is the extent to which trainees can perform a range of desired exploration skills (e.g., reflection of feelings, restatement, open question, talking less) in nonscripted helping sessions after training.

In addition to exploration skills, helpee perception of helper empathy is an important outcome of training. From the psychotherapy literature, we know that client perceptions of empathy are the more highly related to outcome than are other perspec-

tives (see Bohart, Elliott, Greenberg, & Watson, 2002). Thus, it makes more sense to examine helpee perceptions of helper empathy than to rely on judges' ratings, as has been the standard in previous research.

Another possible outcome of helping skills training is helper credibility, which reflects trainees taking on a professional role. Strong (1968), in his social influence theory, posited a two-stage process of interpersonal influence. In the first stage, the therapist gains influence by enhancing his or her perceived credibility. In the second stage, the therapist makes use of the influence he or she has built in the first stage to bring about the desired changes in client behavior and cognitive framework. The importance of credibility and the social influence model have been well documented (see review by Heppner & Claiborn, 1989).

Number of words spoken during helping sessions is another possible outcome. In friendship relationships, people often talk equal amounts. Thus, for people first learning helping skills, one of the big shifts is learning to talk less and correspondingly listen more.

Another important outcome of helping skills training is trainees' self-efficacy for using the helping skills (i.e., feeling more confidence about their ability to use the skills). Self-efficacy may be an interesting outcome variable to examine in its own right, as it reflects growth in confidence as a helper. Also, given the moderate relationship between self-efficacy and performance in the therapist role (see Lent, Hill, & Hoffman, 2003; Lent et al., 2006), assessing trainees' self-efficacy for using the helping skills as an outcome index may yield important information about trainees' ability to utilize the skills in future helping situations effectively.

Finally, rather than just examining pre-post change in self-efficacy, it is important to look at how change occurs throughout the course of training. As instructors, we have noticed that many students start with relatively high confidence because they have been natural helpers, then decline in confidence once they start practicing the skills and learn how difficult they are, and then increase in confidence as they master the skills.

Perspective

Gathering data from multiple perspectives (e.g., trainee, helpee) and types of measures (e.g., self-report, behavioral observations) is important given evidence from the psychotherapy literature (see Hill & Lambert, 2004) that results differ on the basis of perspective and types of measures. Trainees can provide a subjective perspective but may not be able to estimate their skill performance relative to other students. In contrast, clients can provide a perspective on the basis of being recipients of the skills, but they may not be paying much attention to what the helper is doing if they are focusing on their own problems. In addition, there is an advantage for self-report measures in that they reflect the participants' perspective, but the disadvantage is that they require self-awareness. In contrast, behavioral measures (e.g., behavioral counts of the number of skills used in helping session and number of words spoken) provide an "objective" perspective on trainee skill performance but are limited in that they do not capture the experiences of helpers or helpees in sessions. Hence, it seems important to include both trainee and helpee perspectives as well as both subjective and

objective perspectives to gain a more complete assessment of trainee performance of exploration skills.

Predictors of Training Outcomes

In their review, Hill and Lent (2006) found only five studies that investigated predictors of outcome of skills training, each of which investigated a different predictor (i.e., dominance, gender, conceptual level, positive attitudes toward the target skill, and pre-training expectations for nondirective vs. directive therapy style). Therefore, no conclusions can be drawn about predictors because results have not been replicated across studies. Hence, for the present study, we theorized about several potential predictors of the outcomes of helping skills training.

First, we speculated that the trainees who were initially more empathic would have the best training outcomes. Given that training involves teaching students to be empathic and use helping skills (e.g., reflections of feelings), we thought that empathic trainees would have an easier time incorporating the helping skills into their repertoires, and would thus be more effective helpers after training. We further speculated that empathy would have an especially big impact on ability to learn exploration skills (e.g., reflections of feeling, restatements), given that the theoretical basis for exploration skills is Rogers' (1957) person-centered approach, in which empathy is viewed as one of the "necessary and sufficient" conditions of therapy.

Empathy can be conceptualized (Davis, 1980a, 1980b, 1983) as both empathic concern (being warm and compassionate) and as perspective taking (being able to see things from another's viewpoint). The first is more of an emotional connection, whereas the second is more intellectual. Given that these could have different effects, we consider them separately.

Moreover, grade point average (GPA) could also be hypothesized to be a predictor of training outcomes given that grades are often used in academic settings to select and evaluate students. We were unsure, however, of what the relation to training outcome would be. Perhaps students with higher GPAs at the start of the class would be more diligent and thus work harder to learn the skills, even if for purely academic reasons (since grade in the class is at stake). However, some have argued that therapeutic ability is unrelated to intellectual ability (e.g., Carkhuff, 1969). There are times that trainees can clearly grasp the content of what clients are saying but cannot communicate empathically with clients. So grades and academic performance may not be related to being able to intervene effectively with clients.

Finally, we expected that some personality attributes might be negatively related to training outcomes. Specifically, we speculated that perfectionism could hamper a trainee's ability to learn and perform the skills. There are rarely "right" or "perfect" skills to be used in a given moment of a helping session, and this might be frustrating to those trainees who are overly perfectionistic. The findings of a qualitative study by Hill, Sullivan, Knox, and Schlosser (2007) support such a hypothesis, as some of the trainees in this study reported (via journals) that their perfectionistic strivings hindered their development as therapists.

The Present Study

The first purpose of the present study, then, was to investigate the outcomes of helping skills training. First, we explored the

ability of trainees to use exploration skills in helping sessions because these skills were the major focus of the class. We investigated ability to use exploration skills from the perspective of both helpers and clients in the session as well as by behavioral counts of the exploration skills used in the sessions. Second, we investigated helpee perceptions of helper empathy because empathy is a key component of the helping skills model. Third, we investigated the number of words helpers used in helping sessions because listening often involves talking less. Fourth, we examined helpee perceptions of helper credibility because we expected increases in credibility as helpers became more proficient in using the skills. Fifth, we examined changes in self-efficacy for using all the helping skills, given that increases in self-efficacy have been reported consistently for graduate trainees (see Lent et al., 2003). Finally, we examined weekly ratings of self-confidence in using the helping skills to give us an estimate of the process of change during different parts of the semester. We predicted that there would be increases in the use of exploration skills, empathy, credibility, self-efficacy, and confidence, along with decreases in the number of words spoken.

We note that changes in self-efficacy for using the helping skills were assessed using retrospective changes in self-efficacy. Howard (1980) and Bray, Maxwell, and Howard (1984) found that retrospective measures are valid measures of pre-post change, especially when participants have a difficult time estimating themselves on the construct at the beginning. Indeed, pilot data showed that participants were unable to guess at their self-efficacy prior to training; we speculated that they did not know the definitions of the skills, nor did they have an accurate perception of how well they were able to use the skills before training.

We also note that classmates served as the helpees for the trainees/helpers. Given our negative experiences in the past with some undergraduate trainees who were not very skillful before training, we felt that it was not ethical for trainees/helpers to conduct such sessions with recruited clients before any training. In fact, some trainees/helpers are not at all helpful before training and may in fact have an adverse impact on recruited helpees due to a judgmental or advice-giving stance. In addition, we paired students with different classmates for the first and second sessions because dropouts from the class precluded us from having the same persons serve at both points (although we made sure that helpers did not know their partners well). One possible limitation of this practice of using classmates as helping partners is that helpees are learning the skills along with the helpers and are thus not naïve judges of the helper's ability, although it is unclear in which direction judgments might be affected (i.e., this exposure could cause them to be either more or less judgmental than a volunteer client). This methodology does, however, provide an estimate of the helper's ability to conduct a helping session and avoids the social desirability inherent in having the helpee feel pressured to give higher ratings if paired with the same helper for the second session.

A second purpose was to determine whether we could predict what students would demonstrate the best levels of skills after training. Hence, we looked for relationships between four predictor variables (GPA, empathic concern, perspective taking, and perfectionism) and nine training outcomes (trainee/helper self-

report of the use of exploration skills in a session, trainee/helper self-report of the perceived quality of a session, helpee report of the helper use of exploration skills in a session, helpee report of the perceived quality of a session, helpee report of helper empathy in a session, helpee report of helper credibility in a session, behavioral estimate of the proportion of exploration skills used in a session, behavioral estimate of the proportion of words spoken in a session, trainee perceptions of overall self-efficacy for using the helping skills).

For the present study, we investigated the outcomes of helping skills training for undergraduates. Although much of the early research was conducted on graduate trainees (see Baker et al., 1990), training often begins at the undergraduate level. We would expect somewhat different results for undergraduate trainees compared with graduate trainees because undergraduates take helping skills courses for a wider variety of reasons (e.g., to become a therapist to fulfill a credit requirement, to prepare to be a medical doctor or lawyer) than do graduate students (e.g., to become a therapist).

In addition, for the present study, we focused on the Hill (2004) helping skills model. Although most of the previous research was conducted on Carkhuff's (1969) human relations training, Ivey's (1971) microcounseling, and Kagan's (1984) interpersonal process recall, helping skills training has evolved (e.g., it is typically longer and focuses on presenting skills in the context of a helping relationship), and so we need research on helping skills training as it is presently practiced. In the Hill model, which integrates aspects of the earlier Carkhuff (1969), Ivey (1971), and Kagan (1984) models, trainees learn about exploration (e.g., open question, re-statement, reflection of feeling), insight (e.g., challenge, interpretation), and action (e.g., information, direct guidance) skills and then practice them with a great deal of monitoring and feedback. In addition to learning the skills, there is an emphasis on empathy, responsiveness to helpee needs, clinical intuition, and self-awareness. Furthermore, given that only one study (Hill & Kellems, 2002) has investigated the Hill model thus far, more research is needed to determine the effectiveness of training using this model.

Method

Design

This was an exploratory naturalistic study of three undergraduate helping skills classes using Hill's (2004) helping skills model. This study took place at a large mid-Atlantic university with three different instructors each teaching a class of approximately 30 upper-level undergraduate students. Students completed predictor measures at the beginning of the 15-week semester, recorded confidence data weekly after lab sessions in which they practiced the skills, completed measures (both as helpers and helpees) after brief helping sessions at the beginning of the semester and two thirds of the way through the semester, and completed measures (both as helpers and helpees) after a brief helping session at the beginning of the semester and another brief helping session two thirds of the way through the semester. In addition, at the end of the semester, students completed two versions of a measure of self-efficacy (one for present self-perceptions and one for how they perceived that they were prior to training—a retrospective pre).

Helping Skills Class

The class was a 4-credit laboratory class that lasted a 15-week semester. All psychology majors were required to take two laboratory classes, and this class was one of five that fulfilled that requirement. One component of the class (1 hr per week) involved a lecture/discussion of a particular skill. A second hour per week focused on research involving helping skills so that students could learn to be competent consumers of helping skills research. The third component involved a 2-hr weekly laboratory section, in which students practiced each skill in small groups. In addition, each student conducted 20-min helping sessions with a different classmate at the beginning and two thirds of the way through the semester during the lab time. They were required to transcribe the tapes of the two sessions and code each helping skill used; it was stressed that they would be evaluated on how they coded the helping skills rather than on which skills were used. For the labs and sessions, students were told several times throughout the semester that, in accordance with ethical guidelines, they were not required to disclose personal information and indeed could either make up problems or talk about real problems if they wished. At the end of the semester, students voluntarily reported during classroom discussions that they had talked about real problems (mostly roommate concerns, relationship difficulties, academic stresses, career plans), and our reading of the transcripts suggested that real problems were discussed, but we have no firm data about whether indeed everyone talked about real problems. Each class required several written assignments: a self-reflection paper at the beginning of the semester, critiques of research studies, and a research paper involving an analysis of the transcripts of the two helping sessions. Finally, students in each class had to take a midterm and final examination involving multiple-choice, short-answer, and essay questions.

The three instructors met before the beginning of the semester and discussed their syllabi so that they covered similar topics in a similar sequence and agreed on similar policies regarding assignments. They tried to make the courses as comparable as possible.

Participants

The 3 instructors were all European American (Jewish) women. One instructor was a 42-year-old counseling psychologist with a doctorate degree; she had been a teaching assistant for the helping skills class three times and taught the class six times prior to the study. The other 2 instructors were 26- and 27-year-old advanced doctoral students in counseling psychology; they had been teaching assistants for the helping skills class 4–5 times, and had each taught the class once prior to the study. In addition, 8 (5 women, 3 men; 6 European American, 1 Asian American, 1 biracial; 2 seniors, 2 postbaccalaureate, and 4 doctoral students) individuals who had taken helping skills previously served as graduate teaching assistants or lab leaders.

Eighty-five upper-level undergraduate students (74 women, 11 men; 56 European American, 7 African American, 8 Asian, 3 Latino/a, 6 biracial, 5 "other"; 80 seniors, 5 juniors; all psychology majors; mean age was 21.64 years [$SD = 2.31$]) in three helping skills classes taught at the same university during the same semester and participated in the study. We refer to

them as *students* or *trainees* when we discuss training, *helpers* when we discuss their performance in helping sessions, and *helpees* when they served as volunteer clients. Their participation was voluntary and anonymous (i.e., participation in the study was not a requirement of the course).

Predictor Measures

All participants were asked to indicate age, gender, and race/ethnicity on a demographic form. Undergraduate students were also asked to indicate GPA.

The Interpersonal Reactivity Index (IRI; Davis, 1980b, 1983) assesses self-reported empathy. Two subscales of the IRI were used: Empathic Concern (EC), which measures one's level of warmth, compassion, and concern for those who are distressed (e.g., "I often have tender, concerned feelings for people less fortunate than me"), and Perspective Taking (PT), which reflects one's ability to take another's perspective (e.g., "I sometimes try to understand my friends better by imagining how things look from their perspective"). Each subscale contains 7 items (14 total) using a 5-point Likert scale ranging from 0 (*does not describe me well*) to 5 (*describes me very well*). Subscale scores are obtained by summing the items, with higher scores indicating higher levels of empathy. Davis (1983) found a moderate correlation between EC and PT (.33). Davis (1980a) reported that subscales were significantly related to other empathy scales, providing evidence of concurrent validity. Internal consistency alpha was .70–.72 for EC and .78 for PT for Davis (1980a), and .78 for EC and .82 for PT for the present study.

The Almost Perfect Scale-Revised (APS; Slaney, Rice, Mobley, Trippi, & Ashby, 2001) assesses levels of perfectionism. The APS has 23 items (e.g., "I rarely live up to my high standards," "I expect the best from myself") in three subscales (High Standards, Order, and Discrepancy); a 7-point Likert scale ranging from 1 (*strongly disagree*) to 7 (*strongly agree*) is used. Slaney et al. (2001) reported evidence of validity (factor analyses, correlations with other measures of perfectionism, expected correlations with theoretically related constructs) and internal consistency estimates ranging from .85 to .92. In the present study, subscales were highly intercorrelated (from .56 to .84), and alpha for all 23 items was .88, so the total score was used.

Measures Used to Assess Helping Ability During Sessions

The Helping Skills Measure-Exploration (HSM-Exp; Hill & Kellems, 2002) was designed to measure helpee and helper perceptions of helper use of exploration skills. The HSM-Exp has four items scored on a 5-point scale ranging from 1 (*strongly disagree*) to 5 (*strongly agree*). Items are averaged (after reversing negatively worded items) such that higher scores indicate more frequent use of the skills. Exploratory and confirmatory factor analyses of the full helpee-rated HSM indicated three subscales (Exploration, Insight, and Action). An example item is "In this session, my helper asked questions to help me explore what I was thinking or feeling." Validity has been demonstrated by relationships with similar measures and predicted changes across helping skills training. The internal consistency alpha for helpees in Hill and Kellems (2002) was .72, and for the present study it was .71

and .65 (first session) and .61 and .70 (second session) for helpees and helpers, respectively.

The Session Evaluation Scale (SES; Hill & Kellems, 2002) assesses helpee and helper perceptions of the quality of the session. The SES includes four items (e.g., "I thought this session was helpful"), rated on 5-point scales ranging from 1 (*strongly disagree*) to 5 (*strongly agree*). Items are averaged (after reversing negatively worded items) such that higher scores reflect better session outcome. Exploratory and confirmatory principal-axis factor analyses of the helpee version revealed a single factor accounting for 77% of the variance; internal consistency was .91. Internal consistency alpha for Hill and Kellems (2002) was .91, and was .93 and .86 (first session) and .83 and .85 (second session) for helpees and helpers, respectively, in the present study.

The Empathy Scale (ES; Persons & Burns, 1985) assesses client-perceived helper warmth, genuineness, and empathy. The 10 items (e.g., "My therapist understood what I said in today's session") are rated on a 4-point Likert scale ranging from 1 (*not at all*) to 4 (*a lot*). Persons and Burns (1985) found that the ES was positively correlated with clients' change in mood, providing some evidence for validity. The internal consistency alpha was .70 for Persons and Burns and .74 (first session) and .61 (second session) in the present study.

The Counselor Effectiveness Rating Scale (CERS; Atkinson & Carskaddon, 1975) assesses client-perceived helper credibility (i.e., expertness, attractiveness, trustworthiness, and utility). The 10 items (e.g., "the counselor's expertness") are rated on a 7-point Likert scale ranging from 1 (*bad*) to 7 (*good*), and scores are averaged. Atkinson and Wampold (1982) reported a correlation of .80 between the CERS and the Counselor Rating Form (CRF; Barak & LaCrosse, 1975), which provided evidence of concurrent validity. Internal consistency alpha was .90 for Atkinson and Carskaddon (1975) and .87 (first session) and .90 (second session) in the present study.

Behavioral estimates of exploration skills. Skills used in helping sessions were coded using the Helping Skills System (HSS; Hill, 2004), a revision of the Hill Counselor Verbal Response Category System (HCVRCS; Hill, 1978, 1986). The HCVRCS and HSS consist of nominal, mutually exclusive categories of therapist verbal behavior. For the HCVRCS, Hill (1978) established content validity through combining categories from existing measures and having expert therapists from different theoretical orientations determine the representativeness of categories; concurrent validity was established through significant positive correlations between similar categories in other response mode systems (Elliott et al., 1987). For the HSS, Hess, Knox, and Hill (2006) reported an average kappa between pairs of judges of .91. For the present study, each student coded the response modes in each response unit (i.e., grammatical sentence), and then the instructor or teaching assistant corrected the codings. The proportion of exploration skills was calculated by summing the number of open questions, restatements, and reflections of feelings, and dividing by the total number of skills used in the session.

Number of words used. Counts were made of the number of words spoken by helpers and helpees in helping sessions. The number of helper words was divided by the total number of words spoken by both helper and helpee to yield an index of proportion of helper words.

Measures Used to Assess Self-Efficacy and Confidence

The Counseling Activity Self-Efficacy Scales-Helping Skills (CASES-HS; Lent et al., 2003) assesses counselor self-efficacy for performing helping skills. There are three subscales: Exploration Skills, (6 items, $\alpha = .79$), Insight Skills (5 items, $\alpha = .85$), and Action Skills (4 items, $\alpha = .83$). Counselors rate their confidence in their ability to perform each skill on a 10-point scale, ranging from 0 (*no confidence*) to 9 (*complete confidence*). Item responses are averaged, with higher scores reflecting higher self-efficacy. In terms of validity, CASES scale scores related positively and significantly to another measure of counselor self-efficacy, related minimally to a measure of social desirability, evidenced significant change for students over the course of a semester-long practicum, and discriminated among therapists at different levels of experience. We administered the CASES-HS at the end of the semester, asking students to respond for "now" and also retrospectively for before the semester. In essence, then, the CASES-HS represents the trainees' perceptions about how they changed over the semester. Because the three helping skills scales were moderately to highly correlated at both retrospective pre- (r ranged from .58 to .74) and posttesting (r ranged from .47 to .56) and because this measure was administered at the end of the semester after all the helping skills had been covered, the scales were averaged to form a total helping skills scale. Internal consistency alpha was .92 for the retrospective pre- and .88 for the posttest for the present study.

Confidence was assessed via an item that asked, "How confident are you that you could use the helping skills effectively with most clients over the next week?" Note that a single item assessing confidence (using the same format as that used for the CASES; Lent et al., 2003) was used rather than the longer CASES to increase participant cooperation given that this item was administered weekly. At the end of each lab session, participants rated confidence on a 10-point scale ranging from 0 (*no confidence at all*) to 9 (*complete confidence*). Because two of the classes had 11 lab sessions, whereas the third class had 13 lab sessions, the 2 weeks for the third class that did not match the times for the other classes were dropped. Because some students missed lab sessions (31 missed one session, 12 missed two sessions, 1 missed three sessions), the missing data were imputed (a typical approach to missing data in clinical trials, Schafer & Graham, 2002). To impute the data, the expectation maximization program on SPSS was used (this program considers both the mean for an individual participant as well as the data across time for all participants), and thus data for all students on all 11 weeks were included.

Procedures

Prior to the semester, all instructors, teaching assistants, and lab leaders signed a consent form and completed a demographic form.

During the first or second day of each class, the instructor explained the study and encouraged students to participate, but stressed that participation was completely voluntary. The instructors explained that completion of all the measures and helping sessions was required for the class but that each student had the option as to whether to contribute his or her data to the

study for extra credit for the course. Students were assigned code numbers to protect confidentiality. They indicated their code numbers and names on a separate sheet so that their data could be linked together at the end; code sheets were shredded after all the data was entered by undergraduate research assistants who had no instructional connection with any of the three classes.

During the first lab, each trainee was paired with a classmate (someone they did not know), so that each did a 20-min session as a helper and a 20-min session as a helpee (i.e., switched roles). Helpers were instructed to be as helpful as possible in the session but were not given any expectations about what to do. After each session, trainees/helpers completed the HSM-Exp and SES in a random order; helpees completed the HSM-Exp, SES, ES, and CERS in a random order.

During the second lab, students completed the IRI and APS in a random order. To make the completion of measures a learning experience, students were required to write a 5- to 10-page paper describing what they learned about themselves from completing these measures (about cultural background, strengths and weaknesses as a helper, and biases they might bring to a helping situation), although this paper was not used for the present study.

At the end of each weekly lab where they practiced the skills, students rated their confidence in being able to use the helping skills effectively with most helpees during the upcoming week.

About two thirds of the way through the semester (after completing instruction and practice on the exploration stage), trainees conducted another 20-min session with a different classmate (someone they did not know and who had not been in the same small lab group; it was not possible to pair students with the same partners as at the beginning of the semester because of dropouts from the class and scheduling conflicts). Helpers were again instructed to be as helpful as possible in the session but were not given any expectations about what to do. Helpers and helpees again completed the postsession measures in a random order.

Students transcribed the audiotapes of both sessions. They coded their helping skills (checked for accuracy by the instructor or teaching assistant) for each unit (grammatical sentence) in the transcripts; the proportion of exploration skills used was calculated. They also counted (and we recounted for accuracy) the number of words spoken by each person; we calculated the proportion of helper words. Students were required to hand in tapes of sessions to encourage accurate transcription. Instructors erased all tapes to preserve confidentiality.

During the last lab class of the semester, trainees completed the CASES (current and retrospective pretest).

In total, 101 students started the semester in the three classes (ranging from 31 to 36 students per class). One student did not consent to have data used in the study; 2 students dropped out of the class after taking the initial measures; and 13 finished the course but did not provide complete data. Hence, 85 students provided data for this study (24 from one class, 30 from the second, and 31 from the third).

Results

We set alpha at .05 for all analyses. Given the exploratory nature of the present study, our choice of alpha was guided by a desire for

Table 1
Means, Standard Deviations, and Correlations Among Outcome Measures Administered at the Beginning of the Semester

Variable	M	SD	1	2	3	4	5	6	7	8	9
1. CASES-HS	4.57	.90									
2. HSM-Exp-He	4.24	.60	.09								
3. SES-He	4.12	.72	.13	.71***							
4. ES-He	2.71	.31	-.01	.48***	.57***						
5. CERS-He	5.97	.81	.06	.63***	.67***	.67***					
6. HSM-Exp-Hr	4.14	.51	.38**	.21	.24*	.25*	.31**				
7. SES-Hr	3.63	.68	.17	.29**	.36**	.33**	.32**	.52***			
8. % Hr words	0.28	.11	-.02	.06	.00	.02	.08	-.19	.00		
9. % Hr Ex skills	0.28	.15	-.03	.00	-.08	-.18	-.19	.08	.11	-.26*	—

Note. N = 85. CASES-HS = Counseling Activity Self-Efficacy Scales—Helping Skills; HSM-Exp-He = helpee-rated Helping Skills Measure—Exploration; SES-He = helpee-rated Session Evaluation Scale; ES-He = helpee-rated Empathy Scale; CERS-He = helpee-rated Counselor Effectiveness Rating Scale; HSM-Exp-Hr = helper-rated Helping Skills Measure—Exploration; SES-Hr = helper-rated Session Evaluation Scale; % Hr words = proportion of words spoken by helper; % Exp skills = proportion of exploration skills used by the helper (sum of open question, restatements, and reflections of feelings divided by the total number of skills used).
*p < .05. **p < .01. ***p < .001.

a balance between a Type II error (not finding something that might be there) and a Type I error (finding something that might not be accurate).

Preliminary Analyses

Table 1 shows the means and standard deviations of, as well as the correlations among, the outcome measures administered at the beginning of the semester. Table 2 shows the means and standard deviations of, as well as the correlations among, the predictor variables administered at the beginning of the semester and the outcome measures administered at the middle/end of the semester.

If just one item on a short scale (< 5 items) or one or two items on a longer scale (> 5 items) were missing, we filled in the data by inserting the mean for that scale. However, if more items were missing or if the entire scale was not completed, then those

participants were dropped (n = 15). A multivariate analysis of variance (MANOVA), with the four predictor measures (IRI-EC, IRI-PT, GPA, APS) as the dependent variables and the 85 with complete data versus 15 with incomplete data as the independent variable, was not significant using Pillai’s trace, F(7, 91) = 0.43, p > .05, suggesting that the two subgroups did not differ prior to training (note that there was no missing data on the predictor measures). We could not compare the two subgroups on the outcome variables because of missing data on outcome measures for the participants with incomplete data. None of the trainees who had incomplete data were included in subsequent analyses.

Weekly Changes in Confidence

Figure 1 shows the weekly confidence ratings across the 11 lab sessions. A repeated measures analysis of variance (ANOVA)

Table 2
Means, Standard Deviations, and Correlations Among Predictor and Middle-End-of-Semester Outcome Measures

Variable	M	SD	1	2	3	4	5	6	7	8	9	10	11	12	13
1. IRI-EC	4.08	.58													
2. IRI-PT	3.66	.68	.50***												
3. GPA	3.45	.34	-.03	-.18											
4. APS	4.43	.70	.14	.08	-.19										
5. CASES-HS	6.48	.96	.18	.30**	-.21	.04									
6. HSM-E-He	4.56	.44	.23	.11	.03	.08	.03								
7. SES-He	4.40	.51	.07	.07	.27*	.11	-.08	.53***							
8. ES-He	2.80	.22	.06	.07	.09	-.01	.09	.56***	.48***						
9. CERS-He	6.20	.80	.01	.13	.28**	.10	.11	.63***	.60***	.44***					
10. HSM-Exp-Hr	4.40	.48	.15	.21	.07	.07	.27*	.13	.28*	.23*	.29**				
11. SES-Hr	3.98	.53	.08	.09	-.01	.06	.28**	.28*	.36**	.32**	.34***	.50***			
12. % Hr words	0.20	.10	-.16	-.06	-.14	.21	-.15	-.13	-.22*	-.04	-.09	-.23	-.31**		
13. % Hr Ex skills	0.58	.20	.00	-.04	.24*	-.11	-.11	.08	.22*	.12	.20	.31**	.10	-.23*	—

Note. N = 85. IRI-EC = Interpersonal Reactivity Index—Empathic Concern; IRI-PT = Interpersonal Reactivity Index—Perspective Taking; GPA = grade-point average; APS = Almost Perfect Scale—Revised—total scale; CASES-HS = Counseling Activity Self-Efficacy Scales—Helping Skills; HSM-Exp-He = helpee-rated Helping Skills Measure—Exploration; SES-He = helpee-rated Session Evaluation Scale; ES-He = helpee-rated Empathy Scale; CERS-He = helpee-rated Counselor Effectiveness Rating Scale; HSM-Exp-Hr = helper-rated Helping Skills Measure—Exploration; SES-Hr = helper-rated Session Evaluation Scale; % Hr words = proportion of words spoken by helper; % Exp skills = proportion of exploration skills used by the helper (sum of open question, restatements, and reflections of feelings divided by the total number of skills used).
*p < .05. **p < .01. ***p < .001.

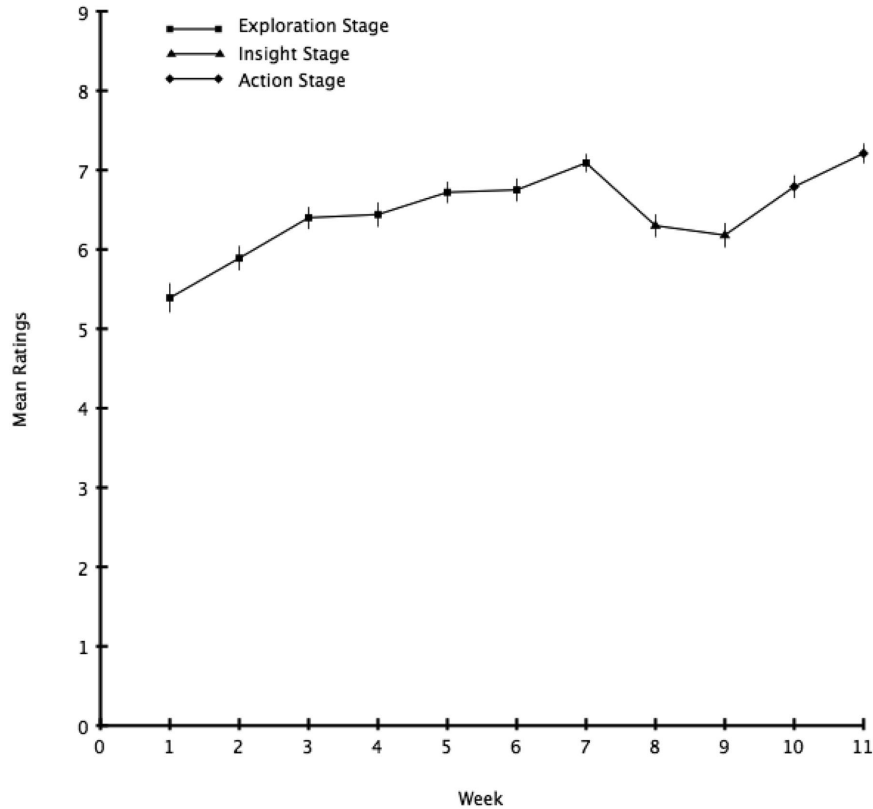


Figure 1. Means and standard errors of weekly confidence measures.

showed significant linear, $F(1, 84) = 58.89, p < .001$; quadratic, $F(1, 84) = 25.05, p < .001$; and cubic, $F(1, 84) = 37.36, p < .001$, effects. Visual inspection of the figure indicated three phases: the first 7 weeks with a linear increase, a decrease in Weeks 8 and 9, and then an increase in Weeks 10 and 11. Post hoc analyses confirmed these findings, in that a repeated measures ANOVA on the first 7 weeks showed just a linear effect, $F(1, 84) = 128.43, p < .001$, whereas a repeated measures ANOVA for the average of Weeks 1–7 ($M = 6.39, SD = 0.97$), Weeks 8 and 9 ($M = 6.24, SD = 1.18$), and Weeks 10 and 11 ($M = 7.00, SD = 1.07$) showed both linear, $F(1, 84) = 36.47, p < .001$, and quadratic effects, $F(1, 84) = 22.96, p < .001$. Hence, these results suggest that trainees increased in confidence across the course of training but had a dip in confidence during Weeks 8 and 9.

A comparison of the findings with the course content on syllabi indicate that students gained steadily in confidence while learning the exploration skills (Weeks 1–7), declined when they started learning insight skills (Weeks 8–9), and then increased again when they started learning action skills (Weeks 10–11). When the sample was divided at the median for both retrospective pre- and post-self-efficacy (CASES-HS) scores, the pattern was the same for both groups, although those with higher self-efficacy on the CASES-HS had consistently higher weekly confidence scores than did those who had lower scores.

Changes Across Training

We analyzed changes across time with repeated measures multivariate analysis of variance (MANOVA) for four sets of

data (helper postsession measures, helpee postsession measures, behavioral session data, and helper-reported self-efficacy). Time was the repeated measure, class was a between-subjects control variable, and the outcome measures were the within-subjects variables. Effect sizes were determined by dividing the differences between the means by the averaged standard deviations; according to Cohen (1988), small = .20–.49, medium = .50–.79, large = .80+.

For helper-rated postsession reports (HSM-Exp and SES), the effects of class and the interaction between time and class were not significant. The overall effect of time was significant, $F(2, 81) = 15.38, p < .001$. Post hoc analyses, using a Bonferroni-corrected alpha of .025 (.05/2 comparisons), indicated significant effects for time for both HSM-Exp, $F(1, 82) = 21.89, p < .001$, medium effect size (.53), and SES, $F(1, 82) = 19.08, p < .001$, medium effect size (.58). Thus, helpers thought that they used more exploration skills in the second than the first session and were more satisfied with the second than the first session.

For helpee-rated postsession measures (HSM-Exp, SES, ES, CERS), the effects of class and the interaction between time and class were not significant. The overall effect of time was significant, $F(4, 79) = 4.30, p < .01$. Post hoc analyses, using a Bonferroni-corrected alpha of .0125 (.05/4 comparisons), indicated significant effects for time for HSM-Exp, $F(1, 82) = 17.14, p < .001$, medium effect size (.62); SES, $F(1, 82) = 8.73, p < .01$, small effect size (.46); and ES, $F(1, 82) = 6.57, p = .01$, small effect size (.34). Effects for time for the CERS were not significant, $F(1, 82) = 4.83, p = .03$, although there was a small effect

size (.29). Thus, helpees evaluated helpers as being more effective (using more exploration skills, were more satisfied with sessions, rated helpers as empathic) in the second session as compared with the first session.

For behavioral data from the sessions (proportion of exploration skills actually used by helpers, and proportion of helper words), we found significant results for class, $F(4, 164) = 5.20, p < .001$, and the interaction between time and class, $F(4, 164) = 6.10, p < .001$. The overall effect of time was also significant, $F(2, 164) = 87.58, p < .001$. Post hoc analyses, using a Bonferroni-corrected alpha of .025 (.05/2 comparisons), indicated significant effects for time for both proportion of exploration skills used, $F(1, 82) = 177.25, p < .001$, large effect size (1.71), and proportion of words, $F(1, 82) = 26.29, p < .001$, medium effect size (.76). Thus, after controlling for the effects of class, helpers used more exploration skills and fewer words in the second as compared with the first session.

For self-efficacy for using helping skills (CASES), the effects of class and the interaction between time and class were not significant. The overall effect of time was significant, $F(1, 82) = 248.89, p < .001$, large effect size (2.05). Thus, at the end of the semester, helpers perceived that their self-efficacy for using helping skills had increased.

Prediction of Outcome

We conducted hierarchical regression analyses on the nine outcome measures (i.e., postsession measures for Session 2 and end-of-semester ratings of self-efficacy). In Step 1, we entered class and the initial score of the outcome measure to control for these effects. In Step 2, we added in the four predictor variables (IRI-EC, IRI-PT, GPA, APS).

After controlling for class and initial levels of the outcome variables, the addition of Step 2 was significant for SES-He, $F(4, 78) = 3.09, p < .05$, and CERS, $F(4, 78) = 3.28, p < .05$. In post hoc analyses for SES-He, GPA was a significant predictor ($\beta = .37; t = 3.28, p < .01$); for CERS, both GPA ($\beta = .33; t = 3.07, p < .01$) and IRI-PT were significant predictors ($\beta = .24; t = 2.01, p < .05$). However, when we deleted outliers (one outlier in the CERS analysis, two outliers in the SES-He analysis), results were no longer significant for either outcome variable. Thus, we were not able to predict which students had better outcomes.

Discussion

After training, undergraduate trainees were able to conduct better sessions while interacting in the moment with classmates sitting across from them talking about real concerns. In particular, trainees used more exploration skills in their second helping session with classmates than they did in the first. It is important to note that this finding was replicated across three different perspectives (behavioral counts of use of exploration skills in the session, helper and helpee ratings of how much exploration skills were used), suggesting that it is a robust finding. These results confirm that undergraduate students can learn to use the exploration skills in half of a semester (the second half of the semester was spent on insight and action skills, which are difficult to assess in brief helping sessions).

Similarly, trainees' use of empathy (as rated by helpees) increased over the semester. This increase in trainees' perceived

empathy mirrors the increase in use of exploration skills from the first to the second session, and makes sense given that exploration skills have often been used as a proxy for empathy (see Hill & Lent, 2006).

In addition, students talked less in their second helping sessions than they had in their first sessions. Talking less reflects listening more and focusing more on the helpee and seems to be a major change that trainees make when they first begin to learn about helping. The shift in talk time can be difficult for beginning trainees who are used to talking and sharing the focus in friendship relationships, so talk time may reflect a shift to taking on a more professional role.

Trainees were also perceived by themselves and by their helpees as conducting better second helping sessions compared with initial sessions. Thus, in addition to finding that trainees used more exploration skills and talked less, these data provide more global impressionistic evidence that trainees learned to manage sessions better and were evaluated as more helpful.

Another major change across the semester was the increase in trainees' self-efficacy for using helping skills, which was assessed both in terms of weekly confidence ratings and end-of-the-semester evaluations. Perhaps having a credible framework to understand the helping process as well as gaining proficiency in specific skills that could be used in helping sessions helped trainees feel more confident in their ability to perform the skills.

Trainees did not, however, increase significantly in terms of credibility as perceived by helpees (although the scores were in the right direction). Some possible explanations for this finding are that credibility may be difficult for classmates to judge (especially in a brief helping session); credibility may be more difficult to change than are helping skills; it may take more than a half a semester to change perceived credibility; or more targeted interventions may be needed in training to influence credibility.

On the basis of these results, we assert that trainees indeed learned the helping skills, thus providing additional evidence to the Hill and Kellems (2002) study for the effectiveness of the Hill model of helping skills training. These findings also extend the findings of previous studies in which analogue taped or written stimuli were used to assess increases in exploration skills and empathy (see review in Hill & Lent, 2006) by using more naturalistic helping sessions and a wider range of outcomes.

The dip in weekly confidence ratings that occurred when the trainees started learning insight skills makes sense in that insight is difficult both to teach and to learn. In addition, being able to do the insight stage seems to require a longer relationship with helpees and a more in-depth knowledge of helpee dynamics (i.e., case conceptualization ability). Moreover, beginning trainees often feel anxious about being intrusive, going "too far," and hurting or insulting their classmates when they use insight skills (see Hill, Sullivan, et al., 2007). In addition, it is not surprising that trainees felt more confident again once action skills were introduced because these skills are usually more familiar to trainees, as it is common to use action skills in friendship/family relationships (e.g., giving advice). Indeed, trainees often jump prematurely to using action skills in their initial helping sessions before they have learned exploration skills.

We were not able to predict outcome from GPA, self-reported empathy (empathic concern, perspective taking), or self-reported

perfectionism. Although some predictors (i.e., dominance, gender, conceptual level, positive attitudes toward the target skill, and pretraining expectations for nondirective vs. directive therapy style) have been investigated in the past literature, these findings have not been replicated across studies. It may be that it is possible to train a wide variety of people to use helping skills, and thus we do not need to be overly concerned with selection. It may also be that more performance-based measures would be better predictors of outcome. In addition, we may have had a fairly skewed sample of people who had already self-selected into this class because they were natural helpers (see Stahl & Hill, 2008).

Strengths and Limitations

The present study has several strengths compared with previous studies. First, we used three classes with different instructors, whereas in most of the previous studies, only one class with one instructor was used. Furthermore, we used a range of predictor and outcome variables, whereas in previous studies, only a limited range of variables was used. Moreover, training was longer (60 hr of class time in the semester vs. 10–30 hr in many of the previous studies) than in many of the previous studies and thus is more similar to how training is now done in college classes. We also used a naturalistic setting, which permitted us to sample both helper and helpee reactions/responses and allows for generalization to other classes teaching this helping skills model in university settings.

One limitation of this study, related to its naturalistic classroom setting where the primary purpose was educational (i.e., not designed for research purposes), was that there was not as much control over variables as would have been possible in a laboratory setting. For example, although the instructors used similar syllabi for the course and covered similar topics, what was emphasized in each class depended to some degree on the style of the instructor, the interests of the students, and the dynamics of the group. Also, because the bulk of the data was turned in at the end of the semester, we did not examine data for completeness until after the semester and thus had to drop 13 participants because of incomplete data. Furthermore, we did not check to determine whether transcripts were transcribed verbatim (although in the hopes of obtaining accurate transcriptions, we told students to turn in tapes and also informed them that evaluation was based on how well they could assess and analyze their interventions rather than on which interventions were used).

In addition, because the research was done in a naturalistic setting, participants were not randomly assigned to classes. Furthermore, we did not have a control group because it was not possible to find an equivalent class in which the students would consent to completing measures and participating in helping sessions. From previous experience, we have found it to be difficult to persuade instructors and students of other helping-related classes (e.g., introduction to counseling psychology) to require that their students complete measures at the beginning and end of the semester and complete helping sessions given that these experiences did not relate to their class content. Furthermore, for logistical reasons, it was difficult to find a course at this university that adequately serves as a “control” group, as many students take their counseling-related courses concurrently in their final semester. Hence, because of the complications inherent in obtaining an

adequate placebo control group, we included only students in helping skills classes in our sample, but caution readers to be aware that we cannot make claims that people without any training would not change in similar ways.

Another limitation relates to the self-report measures, which rely on trainees’ level of awareness and understanding of such characteristics as empathy (i.e., some students may not have a good idea of how empathic they are or indeed may not even know what empathy really is). In addition, social desirability is a potential problem (e.g., students may have been biased in their efforts to appear more or less empathic).

Moreover, classmates served as helpees for helpers, as is the typical practice in this course at this university. Although using classmates rather than volunteer clients is an ethically appropriate practice because students were untrained and not always therapeutic, especially at the beginning of the course, results nevertheless cannot be generalized to real-life therapy. Our goal, however, was not to generalize to real-life therapy but only to training. Also, classmates may have been more cooperative in completing measures and evaluated fellow students more favorably in the second session because they were also being evaluated. However, they could have been more critical in the second session because they knew more about what to look for in a session.

Finally, trainees were undergraduates from a large, selective, public university, and instructors were all female counseling psychologists (one postdoctoral, two predoctoral) with experience in teaching this course. Hence, results can only be generalized to similar trainees, trainers, and settings.

Implications

Given that the present study provides some more evidence that helping skills training for undergraduate students is effective, we need to go to the next step and find out more about the effective components of training. Although there is evidence for the effectiveness of instruction, modeling, and feedback (see review in Hill & Lent, 2006), we need to know more about other components of training such as practice and encouragement (see Hill, Stahl, & Roffman, 2007). We also need to know more about how skills are assimilated into one’s repertoire and whether they are maintained over time with or without further training.

It would also be interesting to further investigate the intriguing findings regarding the dip in confidence ratings during instruction about the insight stage. Our experience has been that it is much more difficult to teach insight skills than to teach exploration and action skills, so perhaps different teaching methods or more time is required for teaching this stage.

In addition, we need to investigate other outcomes (case conceptualization ability, changes in attitudes about seeking help, changes in theoretical orientation, self-awareness) and predictors of outcome (e.g., both trait and state anxiety, narcissism, natural helping ability, amount and quality of previous helping experiences, and motivation to learn and use the helping skills). Also, further investigation of instructor effects would be interesting. Perhaps some instructors are more dogmatic about how the skills should be performed, whereas others are more flexible and encouraging of trainees. Might a dogmatic instructor elicit more or less anxiety among students? What might be the impact on trainee self-efficacy? Furthermore, different types of trainers might be

better with different types of trainees. Finally, we suggest that further research is needed to more directly compare undergraduate and graduate training. It would be interesting to know whether different training methods are effective with the two populations.

References

- Aiken, J., Brownell, A., & Iscoe, I. (1974). The training and utilization of paraprofessionals in a college psychological center. *Journal of College Student Personnel, 15*, 480–486.
- Atkinson, D. R., & Carskaddon, G. (1975). A prestigious introduction, psychological jargon, and perceived counselor credibility. *Journal of Counseling Psychology, 22*, 180–186.
- Atkinson, D. R., & Wampold, B. E. (1982). A comparison of the Counselor Rating Form and the Counselor Effectiveness Rating Scale. *Counselor Education and Supervision, 22*, 25–36.
- Baker, S. B., Daniels, T. G., & Greeley, A. T. (1990). Systematic training of graduate-level therapists: Narrative and meta-analytic reviews of three major programs. *The Counseling Psychologist, 18*, 355–421.
- Barak, A., & LaCrosse, M. B. (1975). Multidimensional perception of counselor behavior. *Journal of Counseling Psychology, 22*, 471–476.
- Bohart, A. C., Elliott, R., Greenberg, L. S., & Watson, J. C. (2002). Empathy. In J. C. Norcross (Ed.), *Psychotherapy relationships that work: Therapist contributions and responsiveness to patients* (pp. 89–108). New York: Oxford University Press.
- Bray, J. H., Maxwell, S. E., & Howard, G. S. (1984). Methods of analysis with response-shift bias. *Educational and Psychological Measurement, 44*, 781–804.
- Carkhuff, R. R. (1969). *Human and helping relations* (Vols. 1 & 2). New York: Holt, Rinehart, & Winston.
- Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed.). Hillsdale, NJ: Erlbaum.
- Cowen, E., Gardner, E., & Zax, M. (1967). *Emergent approaches to mental health problems*. New York: Appelton-Century-Crofts.
- Davis, M. H. (1980a). Individual differences in empathy: A multidimensional approach. *Dissertation Abstracts International, 40*(7-B), 3480.
- Davis, M. H. (1980b). A multidimensional approach to individual differences in empathy. *JSAS Catalog of Selected Documents in Psychology, 10*, 85.
- Davis, M. H. (1983). Measuring individual differences in empathy: Evidence for a multidimensional approach. *Journal of Personality and Social Psychology, 44*, 113–126.
- 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.
- Ford, J. D. (1979). Research on training therapists and clinicians. *Review of Educational Research, 49*, 87–130.
- Guernsey, B. (1969). *Psychotherapeutic agents: New roles for nonprofessionals, parents, and teachers*. New York: Holt, Rinehart, & Winston.
- Heppner, P. P., & Claiborn, C. D. (1989). Social influence research in counseling: A review and critique. *Journal of Counseling Psychology, 36*, 365–387.
- Hess, S., Knox, S., & Hill, C. E. (2006). Teaching graduate student trainees how to manage client anger: A comparison of three types of training. *Psychotherapy Research, 16*, 282–292.
- Hill, C. E. (1978). Development of a counselor verbal response category system. *Journal of Counseling Psychology, 25*, 461–468.
- Hill, C. E. (1986). An overview of the Hill counselor and client verbal response modes category systems. In L. S. Greenberg & W. M. Pinsof (Eds.), *The psychotherapeutic process: A research handbook* (pp. 131–160). New York: Guilford Press.
- Hill, C. E. (2004). *Helping skills: Facilitating exploration, insight, and action* (2nd ed.). Washington DC: American Psychological Association.
- Hill, C. E., & Kellems, I. S. (2002). Development and use of the Helping Skills Measure to assess client perceptions of the effects of training and of helping skills in sessions. *Journal of Counseling Psychology, 49*, 264–272.
- Hill, C. E., & Lambert, M. J. (2004). Methodological issues in studying psychotherapy processes and outcomes. In M. J. Lambert (Ed.), *Handbook of psychotherapy and behavior change* (5th ed., pp. 84–136). New York: Wiley.
- Hill, C. E., & Lent, R. W. (2006). A narrative and meta-analytic review of helping skills training: Time to revive a dormant area of inquiry. *Psychotherapy: Theory, Research, Practice, Training, 43*, 154–172.
- Hill, C. E., Stahl, J., & Roffman, M. (2007). Training novice therapists: Helping skills and beyond. *Psychotherapy: Theory, Research, Practice, Training, 44*, 364–370.
- Hill, C. E., Sullivan, C., Knox, S., & Schlosser, L. (2007). Becoming psychotherapists: Experiences of novice trainees in a beginning graduate class. *Psychotherapy: Theory, Research, Practice, and Training, 44*, 434–449.
- Howard, G. S. (1980). Response-shift bias: A problem in evaluating interventions with pre/post self-reports. *Evaluation Review, 4*, 93–106.
- Ivey, A. E. (1971). *Microcounseling: Innovations in interviewing training*. Springfield, IL: Charles C Thomas.
- Kagan, N. (1984). Interpersonal process recall: Basic methods and recent research. In D. Larson (Ed.), *Teaching psychological skills: Models for giving psychology away* (pp. 229–244). Monterey, CA: Brooks/Cole.
- Kasdorf, J., & Gustafson, K. (1978). Research related to microtraining. In A. E. Ivey & J. Authier (Eds.), *Microcounseling: Innovations in interviewing, counseling, psychotherapy, and psychoeducation* (pp. 323–376). Springfield, IL: Charles C Thomas.
- Klevans, D. R., Volz, H. B., & Friedman, R. M. (1981). A comparison of experiential and observational approaches for enhancing the interpersonal communication skills of speech-language pathology students. *Journal of Speech & Hearing Disorders, 46*, 208–213.
- Korn, P. R. (1980). An undergraduate helping skills course: Skill development and career education. *Teaching of Psychology, 7*, 153–156.
- Lent, R. W., Hill, C. E., & Hoffman, M. A. (2003). Development and validation of the Counselor Activity Self-Efficacy Scales. *Journal of Counseling Psychology, 50*, 97–108.
- Lent, R. W., Hoffman, M. A., Hill, C. E., Treistman, D., Mount, M., & Singley, D. (2006). Client-specific counselor self-efficacy in novice counselors: Relation to perceptions of session quality. *Journal of Counseling Psychology, 53*, 453–463.
- Leventhal, A., Berman, A., McCarthy, B., & Wasserman, G. (1976). Peer counseling on the university campus. *Journal of College Student Personnel, 17*, 504–509.
- Matarazzo, R. G. (1971). The systematic study of learning psychotherapeutic skills. In A. E. Bergin & S. L. Garfield (Eds.), *Handbook of psychotherapy and behavior change: An empirical analysis* (pp. 895–924). New York: Wiley.
- Matarazzo, R. G. (1978). Research on the teaching and learning of psychotherapeutic skills. In A. E. Bergin & S. L. Garfield (Eds.), *Handbook of psychotherapy and behavior change: An empirical analysis* (2nd ed., pp. 921–966). New York: Wiley.
- Payne, P. A., & Woudenberg, R. A. (1978). Helping others and helping yourself: An evaluation of two training modules in a college course. *Teaching of Psychology, 5*, 131–134.
- Persons, J. B., & Burns, D. D. (1985). Mechanisms of action of cognitive therapy: The relative contributions of technical and interpersonal interventions. *Cognitive Therapy and Research, 9*, 539–551.
- Rogers, C. R. (1942). *Counseling and psychotherapy*. Boston: Houghton Mifflin.
- Rogers, C. R. (1951). *Client-centered therapy: Its current practice, implications, and theory*. Boston: Houghton Mifflin.

- Rogers, C. R. (1957). The necessary and sufficient conditions of therapeutic personality change. *Journal of Consulting Psychology, 21*, 95-103.
- Russell, R. R., Crimmings, A. M., & Lent, R. W. (1984). Therapist training and supervision: Theory and research. In S. Brown & R. W. Lent (Eds.), *Handbook of counseling psychology* (pp. 625-681). New York: Wiley.
- Schafer, J. L., & Graham, J. W. (2002). Missing data: Our view of the state of the art. *Psychological Methods, 7*, 147-177.
- Slaney, R. B., Rice, K. G., Mobley, M., Trippi, J., & Ashby, J. (2001). The Revised Almost Perfect Scale. *Measurement and Evaluation in Counseling and Development, 34*, 130-145.
- Strong, S. R. (1968). Counseling: An interpersonal influence process. *Journal of Counseling Psychology, 15*, 215-224.
- Truax, C. B., & Carkhuff, R. R. (1967). *Toward effective counseling and psychotherapy: Training and practice*. Hawthorne, NY: Aldine Publishing.
- Volz, H., Klevans, D., Norton, S., & Putens, D. (1978). Interpersonal communications skills of speech-language pathology undergraduates: The effects of training. *Journal of Speech & Hearing Disorders, 43*, 524-541.

Received August 30, 2007

Revision received March 3, 2008

Accepted March 19, 2008 ■



AMERICAN PSYCHOLOGICAL ASSOCIATION SUBSCRIPTION CLAIMS INFORMATION

Today's Date: _____

We provide this form to assist members, institutions, and nonmember individuals with any subscription problems. With the appropriate information we can begin a resolution. If you use the services of an agent, please do **NOT** duplicate claims through them and directly to us. **PLEASE PRINT CLEARLY AND IN INK IF POSSIBLE.**

PRINT FULL NAME OR KEY NAME OF INSTITUTION _____ MEMBER OR CUSTOMER NUMBER (MAY BE FOUND ON ANY PAST ISSUE LABEL) _____

ADDRESS _____ DATE YOUR ORDER WAS MAILED (OR PHONED) _____

____ PREPAID ____ CHECK ____ CHARGE
CHECK/CARD CLEARED DATE: _____

CITY _____ STATE/COUNTRY _____ ZIP _____

(If possible, send a copy, front and back, of your cancelled check to help us in our research of your claim.)

YOUR NAME AND PHONE NUMBER _____ ISSUES: ____ MISSING ____ DAMAGED

TITLE	VOLUME OR YEAR	NUMBER OR MONTH
_____	_____	_____
_____	_____	_____
_____	_____	_____

Thank you. Once a claim is received and resolved, delivery of replacement issues routinely takes 4-6 weeks.

(TO BE FILLED OUT BY APA STAFF)

DATE RECEIVED: _____	DATE OF ACTION: _____
ACTION TAKEN: _____	INV. NO. & DATE: _____
STAFF NAME: _____	LABEL NO. & DATE: _____

Send this form to APA Subscription Claims, 750 First Street, NE, Washington, DC 20002-4242

PLEASE DO NOT REMOVE. A PHOTOCOPY MAY BE USED.