


# Helping Skills Training for Undergraduate Students: Who Should We Select and Train? $\Psi$

The Counseling Psychologist  
2016, Vol. 44(1) 50–77  
© The Author(s) 2015  
Reprints and permissions:  
sagepub.com/journalsPermissions.nav  
DOI: 10.1177/0011000015613142  
tcp.sagepub.com  


**Clara E. Hill<sup>1</sup>, Timothy Anderson<sup>2</sup>, Kathryn Kline<sup>1</sup>,  
Andrew McClintock<sup>2</sup>, Saryn Cranston<sup>2</sup>,  
Shannon McCarrick<sup>2</sup>, Allison Petrarca<sup>2</sup>,  
Lina Himawan<sup>2</sup>, Andrés E. Pérez-Rojas<sup>1</sup>,  
Avantika Bhatia<sup>1</sup>, Shudarshana Gupta<sup>1</sup>,  
and Margo Gregor<sup>1</sup>**

## Abstract

We examined the effectiveness of the Hill model of helping skills training for 191 undergraduate students in six sections of a semester-long course. Students completed self-report, performance, and nonverbal measures at the beginning; they conducted one 20-min helping session at the beginning and another toward the end of the semester; and they completed self-efficacy measures at the end of the semester. Students' helping skills improved over the course of the semester, as evidenced by higher helper- and volunteer client-rated session quality, reduced proportion of words spoken in sessions, increased proportion of exploration skills used in sessions, and increased self-efficacy for using helping skills. Self-reported empathy predicted four of the five helping skills criteria at the beginning-of-semester assessment. Facilitative interpersonal skills predicted end-of-semester self-efficacy in helping skills when controlling for retrospective prelevels and instructor effects. Implications for training and research are presented.

---

<sup>1</sup>University of Maryland, College Park, MD, USA

<sup>2</sup>Ohio University, Athens, OH, USA

## Corresponding Author:

Clara E. Hill, Department of Psychology, University of Maryland, College Park, MD 20742, USA.  
Email: cehill@umd.edu

---

$\Psi$  The Division 17 logo denotes that this article is designated as a CE article. To purchase the CE Test, please visit [www.apa.org/ed/ce](http://www.apa.org/ed/ce)

**Keywords**

helping skills training, empathic personality traits, exploration skills, helper self-efficacy, facilitative interpersonal skills

Several reviewers (Baker & Daniels, 1989; Baker, Daniels, & Greeley, 1990; Ford, 1979; Hill & Knox, 2013; Hill & Lent, 2006; Kasdorf & Gustafson, 1978; Lambert, DeJulio, & Stein, 1978; Matarazzo, 1971, 1978; Matarazzo & Patterson, 1986; Russell, Crimmings, & Lent, 1984) have established the effectiveness of helping skills training using Carkhuff's Human Relations Training (HRT; 1969), Ivey's Microcounseling (MC; 1971), and Kagan's Interpersonal Process Recall (IPR; 1984) models. More recently, two studies have demonstrated the effectiveness of Hill's helping skills model (Hill & Kellems, 2002; Hill et al., 2008), and an additional three studies (Chui et al., 2014; Jackson et al., 2014; Spangler et al., 2014) have shown the effectiveness of teaching specific insight skills within Hill's helping skills model. As Hill and Lent (2006) noted, however, questions remain as to the selection and trainability of trainees. Are some students naturally skilled helpers prior to training? Are some students more likely than others to benefit from training?

The investigation of moderators is particularly important for tailoring training programs to fit students' needs. In addition, graduate programs in the helping professions could benefit if we could identify what types of students would be most likely to learn helping skills. From another perspective, undergraduates could become more informed about their likelihood of success in helping professions through assessment of their helping abilities. Thus, the overall purpose of the present study was to investigate predictors of initial helping ability and outcomes of helping skills training.

**Description of the Hill Helping Skills Model**

Ridley, Kelly, and Mollen (2001) evaluated Hill's helping skills model (Hill, 2004, 2009, 2014; Hill & O'Brien, 1999) as having the best overall coverage compared with other models currently in use in terms of skills, culture, theory, cognition and affect; integration of skills with cognition and affect; and relationship of skills to therapeutic change. The Hill model primarily grew out of Carkhuff's (1969) HRT with some influence of Ivey's (1971) MC and Kagan's (1984) IPR. In addition, the model was modified extensively based on years of teaching the skills to undergraduate and graduate students as well as conducting and incorporating the results of empirical research on therapist techniques and psychotherapy process.

There are specific goals and skills for each of the three stages of the Hill model. The goals of the exploration stage are to build the relationship, help clients tell their stories, and help clients explore thoughts and feelings using open questions, restatements, reflections of feelings, and disclosures of feelings. The goals of the insight stage are to facilitate awareness, promote insight, and process the therapeutic relationship, using challenges, open questions, interpretations, disclosures of insight, and immediacy. The goals of the action stage are to help clients in terms of relaxation, behavior change, behavior rehearsal, and decision making, using open questions, information, and direct guidance. It is important to note that the skills are not to be applied in a rigid, "cookbook manner." Rather, students are taught that a variety of skills can be helpful for each different goal, and that an empathic manner is crucial. Thus, empathy, collaboration, flexibility, clinical intuition, ethics, self-awareness, awareness of client reactions, and case conceptualization are heavily emphasized in training, in addition to building individual skills. Research has shown that instruction, modeling, practice, and feedback are all important components of the training process (Hill, Spangler, Jackson, & Chui, 2014).

## **Moderators of the Hill Model of Helping Skills Training**

We found only one study that investigated potential moderators of helping skills training programs for undergraduate students (although there were some studies of brief training for specific skills). Hill et al. (2008) found that training outcomes (e.g., helper and client ratings of helping skills, overall session quality, changes in self-efficacy) were not predicted by initial grade point average, self-rated empathy, or self-rated perfectionism. A more comprehensive investigation thus seemed warranted, casting a wider net for possible moderators (i.e., predictors).

When selecting possible predictors, we generally considered variables that reflect an empathic style or interpersonal ability, given that the Hill (2004, 2009, 2014) model is based on Rogers's (1957) client-centered approach, which emphasizes the importance of the facilitative conditions (empathy, genuineness, warmth). Thus, the skills of exploration, insight, and action can be considered manifestations of the facilitative conditions if implemented empathically. Hence, we expected that persons who have more interpersonal ability would be more likely to use the skills better prior to training. They may also be able to learn to use the skills more effectively as a result of training. We suggest that interpersonal ability can be measured through (a) self-report measures of helping ability, (b) helping experiences, (c) performance-based measures, and (d) nonverbal measures.

## *Self-Report Measures of Helping Ability*

Although not found to be a significant predictor in Hill et al.'s (2008) study, trait empathy (Davis, 1983), which is one's capacity to cognitively and emotionally understand another's experience, seems like a trait that should be related to helping ability. As postulated by Hill and Lent (2006), students who are empathic (or in this case, who perceive themselves to be empathic) should be more effective helpers and should more readily learn and incorporate helping skills, particularly exploratory helping skills, than would less empathic students.

Similarly, natural helping ability (Stahl & Hill, 2008), which is an untrained ability to be helpful, should predict initial helping ability and the outcomes of training. Natural helping ability might be a good predictor of undergraduate helping ability because it reflects personal qualities that likely correlate highly with the facilitative conditions.

A potential problem with self-report measures is that individuals often are unable to accurately assess themselves, especially on highly subjective traits and abilities such as interpersonal qualities, emotional responsiveness, and facilitative behaviors (Dunning, Heath, & Suls, 2004; Ogles, 2013). Indeed, in academic and workplace settings, the relationship between self-rated ability or knowledge and actual performance is often quite weak (e.g., Chemers, Hu, & Garcia, 2001; Mabe & West, 1982; Stajkovic & Luchins, 1998). Nonetheless, we included self-report measures in the present study because they are face valid and also because we wanted to determine whether they are related to non-self-report measures of interpersonal ability.

## *Helping Experiences*

It makes sense that people with natural talent for helping others would seek out opportunities to help others (e.g., peer counseling, camp counseling) and that such experiences would help to increase their helping ability. Indeed, Jackson et al. (2014) and Spangler et al. (2014) found that prior helping experiences predicted outcomes of training for interpretations and immediacy skills, respectively, although Chui et al. (2014) did not find the prior helping experiences predicted outcomes of training for challenging skills. In an effort to replicate and extend such findings to entire training programs, we included a measure of prior helping experiences, as operationalized by Spangler et al. (2014).

## *Performance Measures*

It seems reasonable to expect that performance in an analog helping situation would predict performance in a helping session. Indeed, situational performance tests are commonly used in industry for selecting employees (McDaniel,

Morgeson, Finnegan, Campion, & Braverman, 2001). One such potential performance measure is the Facilitative Interpersonal Skills (FIS; Anderson, Ogles, Patterson, Lambert, & Vermeersch, 2009) task, which assesses a helper's ability to provide facilitative conditions (e.g., warmth, empathy) and to be actively engaged (e.g., persuasive) in an analog therapy session. In two naturalistic studies, clients of therapists with a higher FIS scores had a greater decrease in symptoms over the course of treatment than did clients of therapists with lower FIS scores (Anderson, McClintock, Himawan, Song, & Patterson, in press; Anderson et al., 2009). Furthermore, in a randomized clinical trial in which therapists were selected for their low and high interpersonal skills (based on the FIS and a social skills measure), therapists who were high on interpersonal skills had a better alliance and outcome than did therapists who were low on interpersonal skills (Anderson, Crowley, Himawan, Holmberg, & Uhlin, 2015).

### *Nonverbal Measures*

Training outcomes might also be predicted by nonverbal measures of emotion recognition (i.e., the ability to interpret emotional expressions in the face, voice, or body posture) given that such measures are more objective than self-report (Gallese, Keysers, & Rizzolatti, 2004). The ability to read nonverbal cues of emotion is particularly relevant to the present study because of the link between nonverbal emotion recognition and interpersonal competency, empathy, emotional intelligence, and social skills (Austin, 2004; Cook & Saucier, 2010; Lawrence, Shaw, Baker, Baron-Cohen, & David, 2004; Padykula & Horwitz, 2012). The Reading of Mind in Eyes Test (RMET; Baron-Cohen, Wheelwright, Hill, Raste, & Plumb, 2001), which assesses ability to identify emotional states based on identifying emotions from photographs of a person's eyes, seems like a promising measure given its good psychometric properties (see Measures section).

### **Purposes of the Present Study**

The first purpose of the present study was to determine whether we could replicate past studies (Hill & Kellems, 2002; Hill et al., 2008) regarding the effectiveness of helping skills training for undergraduate students using the overall Hill (2009, 2014) model. We hypothesized that students' helping skills would improve during a semester-long helping skills training course, as evidenced by higher volunteer client and helper evaluations of the quality of brief helping sessions, the use of more exploration skills and fewer words in brief helping sessions, and perceptions of increased self-efficacy for using helping skills. We describe the nature and structure of the course in more detail in the next section.

Our second purpose was to use multiple methods to predict helping ability prior to training. We postulated that helping ability prior to training would be related to self-reported empathy and natural helping ability, prior helping experiences, performance on a FIS task, and scores on a nonverbal measure of emotional sensitivity.

Our third purpose was to predict helping ability after training. We postulated that helping performance and self-efficacy for using the skills near the end of training would be related to self-reported empathy and natural helping ability, prior helping experiences, performance on a FIS task, and scores on a nonverbal measure of emotional sensitivity, after controlling for performance at the beginning of training.

We note that, following Hill et al. (2008), initial levels of self-efficacy for using helping skills were assessed using retrospective assessments. Howard (1980) and Bray, Maxwell, and Howard (1984) found that retrospective measures are valid measures, especially when participants have a difficult time estimating themselves on the construct at the beginning. Indeed, Hill et al. reported pilot data showing that participants were not able, prior to training, to estimate self-efficacy, which they suggested was because participants did not have accurate perceptions of how well they were able to use the skills before training.

## Method

### *Helping Skills Course*

This 15-week, 4-credit psychology laboratory course focused on teaching the skills in Hill's (2004, 2009, 2014) three-stage (exploration, insight, action) model of helping. The first half of the course focused on exploration skills, whereas the second half focused on insight and action skills. The course entailed a 2-hr weekly lecture/discussion/modeling of helping skills in classes of 30 to 40 students. These classes were divided into smaller group of eight to 12 students for 2-hr weekly laboratory meetings. In these smaller labs, students practiced applying skills in the group or in dyads with students rotating serving as volunteer clients and helpers. Confidentiality and ethics were stressed in the course because of the nature of disclosure in the helping sessions. Students were told that, in accordance with ethical guidelines, they were not required to disclose personal information and could choose to either invent problems or discuss real ones when they were volunteer clients. Although some content may have been fabricated, students generally indicated at the end of the semester that they had discussed real problems (e.g., roommate concerns, relationship difficulties, academic stress, career plans).

To minimize variation among classes, the instructors met before the start of each semester to discuss their syllabi; they planned to cover topics in a similar sequence and agreed upon policies regarding assignments. Considerable variation occurred, however, in terms of how lecture, discussion, modeling, practice, and feedback were implemented based on instructor preference and student characteristics. Students conducted a 20-min helping session with a classmate at the beginning of the semester and another approximately 3 weeks before the end of the semester. They were also required to write a self-reflection paper at the beginning of the semester, to transcribe and code both helping skills sessions, to write a research paper based on their helping sessions, and to take midterm and final examinations.

### Participants

**Instructors.** Four (two female Southeast Asian Indian, one Latino male, one female European American) instructors, all doctoral students in counseling psychology, participated in the study. All had been teaching assistants for the course four to eight times. Two instructors taught a class in each of two semesters, whereas the other two taught only one class each. Instructors reported that they believed in the Hill helping skills model an average of 7.25 ( $SD = 0.96$ ) on a 9-point scale (1 = *not at all*, 9 = *completely*). Using a 5-point scale (1 = *low*, 5 = *high*), instructors reported belief and adherence to the following orientations: psychoanalytic/psychodynamic ( $M = 5.00$ ,  $SD = 0$ ), humanistic ( $M = 4.25$ ,  $SD = 0.50$ ), feminist/multicultural ( $M = 3.25$ ,  $SD = 0.96$ ), and cognitive-behavioral ( $M = 3.00$ ,  $SD = 0.82$ ). All were authors of the present study.

**Students.** Participants were 191 (141 female, 50 male; 115 European American, 28 Asian, 20 multiethnic or Other, 19 African American, nine Latino; one sophomore, 31 juniors, 159 seniors; 190 psychology majors, one other major; age  $M = 21.35$ ,  $SD = 2.00$  years) upper-level undergraduate students in six helping skills classes (39 to 47 students per class) taught at one large public Mid-Atlantic university during two semesters. All students had taken several prerequisite courses (e.g., introductory psychology, statistics), and 107 (56%) had also previously taken introduction to counseling and/or introduction to clinical psychology courses. Participation in the study was voluntary (i.e., not a requirement of the course) and confidential; students obtained extra credit for participating (alternate extra credit experiences were provided).

**Judges.** Eleven students (three male, eight female; 10 European American, one African American) at a large public Midwestern university served as

judges for coding the FIS data. Of these 11, four were clinical psychology doctoral students (with 1 to 4 years of training) who served as co-leaders for coding teams. The other seven were undergraduate students. Judges ranged in age from 20 to 27 years.

### **Predictor Measures**

**Demographic measure.** This form was used to collect the following information: age, gender, race/ethnicity, current year in school, career interests and plans, highest educational degree, major, and previous coursework in counseling/clinical psychology.

**Interpersonal Reactivity Index (IRI).** The IRI (Davis, 1983) is a widely used, self-report measure of empathy. The 28 items on the IRI are rated on a 5-point Likert-type scale, ranging from 0 (*does not describe me well*) to 4 (*does describe me very well*). The IRI contains four subscales, each composed of seven items. The Perspective Taking and Fantasy subscales measure the cognitive aspects of empathy, whereas the Empathic Concern and Personal Distress subscales measure the emotional aspects of empathy. The sum of these subscales provides an overall estimate of trait empathy, with total scores ranging from 0 to 112. Davis (1983) reported that all four subscales were positively correlated with other measures of empathy, suggesting concurrent validity. Davis (1980) reported internal consistency ( $\alpha$ ) estimates ranging from .71 to .77, and test–retest reliability from .62 to .71. Following Stahl and Hill (2008), we used only the items from the three positively worded subscales (excluding Personal Distress); internal consistency  $\alpha$  for the present study was .78.

**Natural Helping Measure (NHM).** The NHM (Stahl & Hill, 2008) assesses an individual's helping inclinations and experiences. The NHM includes five items (e.g., "I often find myself helping others with their problems") that are rated using a 7-point Likert-type scale ranging from 1 (*never*) to 7 (*always*); the final score is the average of the five items. Total scores range from 1 to 7. Stahl and Hill (2008) reported a one-factor structure based on a principal-axis factor analysis, accounting for 51% of the variance with all items loading  $>.50$ . Stahl and Hill found the test–retest reliability over a 2- to 4-week period to be .67; internal consistency alpha was .81. Internal consistency alpha for the present study was .90.

**Prior Helping Experiences Measure (PHE).** This four-item self-report measure (Spangler et al., 2014) was created to assess the amount of experience one



has in providing help to others. Questions assessed level of experience (e.g., “Completed coursework in peer counseling, peer mediation, or helping skills”), practical experience (e.g., “I have had experience helping clients directly since coming to college, in a setting such as the Help Center, as a peer counselor or resident assistant, in a psychiatric care setting, or in case management”), and didactic experience (e.g., “I have read books about counseling”). All items used a 5-point scale, ranging from 0 (*none at all*) to 4 (*extensive*); the final score is the average of the four items. Total scores range from 0 to 4. A factor analysis revealed a single factor accounting for 66% of the variance; all items loaded greater than .59; internal consistency  $\alpha$  was .82 (Spangler et al., 2014). Internal consistency  $\alpha$  for the present study was .81.

**FIS performance task.** In the FIS performance task (Anderson et al., 2009), students are asked to respond to seven brief video clips (approximately 1 min each) depicting challenging scenarios in therapy. Scenarios were chosen from third-session recordings in a well-researched psychotherapy archive (Strupp, 1993) to represent (a) the most challenging situations for therapists, (b) situations in which clients directly referred to troubling aspects of their relationship with their therapist, and (c) a range of client interpersonal styles (e.g., clients being too friendly, hostile, controlling, or submissive). Actors were hired to reenact transcripts of the actual sessions, and trained clinicians coached the actors to capture the interpersonal style of the participants and the tensions within the enacted relationships. Each of these simulated clients was filmed in full view over the shoulder of the therapist. For further details on the creation of the FIS stimuli, see Anderson et al. (2009) and Anderson et al. (in press).

In this study, participants were tested individually. Each participant was instructed to act as if he or she were the therapist for each client. After each of the FIS video clips, the video was paused and text appeared, “It’s your turn to talk.” Participants were given as long as needed to respond, and the next scenario was started when they finished responding. Participant responses were recorded by a video camera which had been placed slightly to the right of where the participant sat.

Each video-recorded response was coded by trained judges (see Procedures section) on (a) warmth, acceptance, and understanding; (b) empathic accuracy; (c) alliance-bond capacity; (d) verbal fluency; (e) emotional expression; (f) persuasiveness; (g) hope and positive expectations; and (h) alliance rupture-repair responsiveness. A 5-point Likert-type scale was used, ranging from 1 (*skill deficits*) to 5 (*optimal presence of the skill*). Judges used a rating of “3” as a baseline and increased or decreased ratings based on evidence of skills present or lacking in the responses. The overall FIS rating was an

average of the eight FIS items. In prior studies (e.g., Anderson et al., 2015; Anderson et al., in press; Anderson et al., 2009), acceptable interrater reliability was found (intraclass correlation [ICC] = .80 to .86), internal consistency among items was high ( $\alpha = .95$ ), and FIS predicted psychotherapy outcomes. Concurrent validity was mixed, in that FIS was positively correlated with self-reported social skills, empathy, and sociability, but not with psychological mindedness (Anderson et al., 2015).

In the present study, interrater reliability among raters within six teams ranged from .91 to .95. Data were averaged across all coders for each item for each participant. The internal consistency for the eight FIS items was  $\alpha = .97$ . Given that a principal components analysis with oblimin rotation yielded one factor accounting for 83% of the variance, a total score was calculated by averaging the eight FIS items.

**RMET.** RMET (Baron-Cohen et al., 2001) assesses the ability of respondents to identify emotional states based on photographs of a person's eyes. The participant is presented with a series of 36 pairs of eyes and asked to choose which of four words most closely matches the mental state conveyed (e.g., playful, jealous). The measure is scored by adding the number of correct responses. Total scores range from 0 to 36. The Kuder Richardson 20 (KR20) reliability estimate for this sample was .65.

## Outcome Measures

**Session Process and Outcome Measures (SPOM).** The SPOM (Hill & Kellems, 2002) was designed to measure client perceptions of helper abilities in sessions. Exploratory and confirmatory factor analyses indicated three scales of the 21-item SPOM: (a) The Helping Skills Measure (HSM) with four- to five-item subscales of Exploration ( $\alpha = .73$ ; for example, "In this session, my helper asked questions to help me explore what I was thinking or feeling"), Insight ( $\alpha = .71$ ; for example, "In this session, my helper helped me understand the reasons behind my thoughts, feelings, and/or behaviors"), and Action ( $\alpha = .82$ ; "In this session, my helper helped me figure out how to solve a specific problem"); (b) the Relationship Scale (RS;  $\alpha = .81$ ; for example, "In this session, I trusted my helper"); and (c) the Session Evaluation Scale (SES;  $\alpha = .88$ ; for example, "I thought this session was helpful"). All items are scored on a 5-point scale (1 = *strongly disagree*, 5 = *strongly agree*). Items are averaged (after reversing negatively worded items) for a total score between one and five, such that higher scores indicate more frequent use of the skills. Validity was demonstrated by relationships with similar measures and prediction of changes across helping skills training. Hill and Kellems

found correlations of .43 to .63 among the five scales for clients; in the present study, the correlations ranged from .45 to .69, so we combined these five scales into a composite score called SPOM-C representing client perceptions of session quality. A parallel form was created for helpers and has been used in several previous studies (e.g., Hill et al., 2008). For helpers, the range of correlations in the present study among the five scales ranged from .31 to .71, so we combined those into a composite score called SPOM-Hr representing helper perceptions of session quality. Correlations between the SPOM-C and SPOM-Hr were moderate (.39). Good internal consistency was found for both the overall SPOM-Hr for both Session 1 ( $\alpha = .90$ ) and Session 2 ( $\alpha = .87$ ), and for SPOM-C for both Session 1 ( $\alpha = .92$ ) and Session 2 ( $\alpha = .87$ ).

*Exploration skills.* Skills used in helping sessions were coded using the Helping Skills System (HSS; Hill, 2009), 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. Furthermore, Hill had 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, as in Hill et al. (2008), each student coded the response modes in each response unit (i.e., grammatical sentence) of the typed transcript, and then the instructor or teaching assistant corrected the codings. The HSS includes 12 nominal, mutually exclusive categories of verbal behavior, organized into three larger categories of exploration skills (including approval and reassurance, closed questions, open questions, restatements, and reflections of feelings), insight skills (includes challenges, interpretations, self-disclosure, and immediacy), and action skills (includes information and direct guidance). The proportion of exploration skills was calculated by dividing the total the number of exploration skills by the total number of skills used in the session.

*Helper words.* We calculated the proportion of helper words by dividing the number of helper words by the total number of words spoken by both helper and client.

*Counseling Activity Self-Efficacy Scales—Helping Skills (CASES-HS).* CASES (Lent, Hill, & Hoffman, 2003) assesses counselor self-efficacy for performing

helping skills. There are three subscales: Exploration Skills (six items,  $\alpha = .79$ ), Insight Skills (five items,  $\alpha = .85$ ), and Action Skills (four items,  $\alpha = .83$ ). Counselors rate their confidence in their ability to perform each skill (e.g., attending) on a 10-point scale, ranging from 0 (*no confidence*) to 9 (*complete confidence*). A total score of the measure was calculated by averaging the item responses into scores ranging from 0 to 9, with higher scores reflecting higher self-efficacy. In terms of validity, CASES scale scores correlated 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 at the end of the semester, asking students to respond for “now” (post) and also retrospectively for “then” or before the semester (pre). These ratings thus represent trainees’ perceptions of change over the semester. For the present study, internal consistency  $\alpha$  was .91 for the CASES-then (retrospective pre) and  $\alpha = .92$  for the CASES-now (post). Correlations among the three subscale scores in Lent et al. (2003) ranged from .52 to .62, and in the present study ranged from .51 to .69, so these were combined into a total score.

## Procedures

We obtained institutional review board (IRB) approval and followed all ethical standards. Prior to the semester, all instructors signed an informed consent form and completed a demographic form. Code numbers were used rather than names on all data to protect confidentiality.

*Recruiting participants and pretraining testing.* During the first or second day of the fall 2013 semester, a recruiter not associated with the specific class asked for volunteers to participate in the study. Students were informed that they would complete measures outside of class for a total time commitment of about 1 to 2 hr and that they would receive extra credit for participation. In total, 119 students started the fall 2013 semester in one of three classes (ranging from 39 to 40 students per class); 23 students did not provide consent, and 64 finished the course but did not provide complete data. Hence, 32 students (27%) provided complete data for this study (eight to 12 per class).

Because the participation rate was low, we were concerned that this sample would not represent the full range of abilities. We thus negotiated with the IRB so that students were required to complete all of the measures as part of the class in the spring 2014 semester, although students could still choose whether to submit their data for research (consistent with Chui et al., 2014; Hill et al., 2008; Jackson et al., 2014; Spangler et al., 2014, studies).

Alternative options for earning extra credit (e.g., other studies) were also provided during both semesters. For the second semester, 138 students started in three classes (ranging from 44 to 47 students per class); one student dropped out of the class during the middle of the semester and 22 students finished the course but did not provide complete data. Hence, 115 students (83%) from the spring semester provided complete data for this study.

During class, after listening to the recruitment script, students were given time to read and sign the informed consent document. They completed several measures (demographic, PHE, NHM, IRI, RMET) outside of class on Qualtrics, a web-based survey software. Students were also scheduled for individual appointments during the first month of the course to complete the FIS task. Each student completed the FIS task in a private room equipped with a computer screen, recorder, and chairs arranged to simulate a helping setting. The video was then started, commencing with a brief video role induction during which the student was instructed to act as if she or he were a therapist in the upcoming client video situations. For each of seven simulated clients, the student listened to approximately 30 s of background information about the client and then watched the 1-min therapy video clip. At the end of each clip, the image of the client's face was frozen with the superimposed words, "It's your turn to talk" appearing on the video screen. Participants were allowed as much time as they needed to respond while being video recorded.

*Beginning-of-semester assessment of helping skills in brief sessions.* During the first lab, each trainee was randomly paired with a classmate whom they did not know for a 20-min session as a helper and another 20-min session as a client (i.e., switched roles). Helpers were instructed to be as helpful as possible in the session but were not given any specific guidelines about what to do. After each session, trainees/helpers and clients completed their respective SPOM forms. Trainees/helpers then transcribed their own sessions using the audiotapes, coded all the helping skills for each unit (grammatical sentence) in the transcripts, and counted the number of words used by both helper and client in the session. Transcripts were checked for accuracy by the instructor or the teaching assistant assigned to the instructor. Audiotapes were erased after the transcripts were completed.

*Late-semester assessment of helping skills in brief sessions.* About two thirds of the way through the semester (after having completed training on the exploration and insight stages), trainees were randomly paired with another classmate for a 20-min session using the same procedures as mentioned previously.

*Assessment of self-efficacy during last week of semester.* Finally, trainees completed the CASES-HS (current and retrospective pretest). Instructors were

informed about who participated after all the grading was done for the classes so that they could assign extra credit.

*Coding the FIS.* Graduate student judges were trained by the second author for 2 months using FIS data from previous studies until high interrater reliability was reached (ICCs > .80). These graduate students then trained undergraduate judges and served as team leaders. Participants' video-recorded responses to the FIS video stimulus clips were rated in random order with rotating teams of four to five judges per case (all responses for a given case were rated by the same team). Each judge independently watched and coded the FIS responses. Weekly meetings across groups were held to discuss differences in ratings, to calibrate judges with the principles of the coding manual, and to reduce drift (although independent judgments were never changed).

### Data Analysis

Analyses of changes in the use of helping skills were conducted using repeated measures analyses of variance (ANOVAs) for the five helping skills dependent variables: CASES-HS, SPOM-C, SPOM-Hr, exploratory skills, and helper words.

The set of five predictor variables (i.e., IRI, NHM, PHE, FIS, and RMET) were then examined for whether they predicted the five helping skills variables. Hierarchical linear models (HLM; Raudenbush & Bryk, 2002; Singer, 1998) were conducted to identify whether the set of five predictors, entered simultaneously, were predictive of each of five helping skill variables at (a) beginning-of-semester levels of helping skills and (b) late-semester helping skills. Both were two-level HLMs, with students nested within instructors (see Results section for additional reasons for including instructor within this model). Because instructors were nested within semesters, we tried to fit a three-level HLM (students nested within instructors, and instructors nested within semesters). The results, however, showed that there was not sufficient variability due to semesters to warrant adding the third level to the HLM, and so the remainder of the analyses used a two-level model.

Thus, beginning-of-semester helping skills was modeled as follows:

$$HS_{ij} = \pi_{00} + \pi_{10} (IRI)_{ij} + \pi_{20} (NHM)_{ij} + \pi_{30} (PH)_{ij} \\ + \pi_{40} (FIS)_{ij} + \pi_{50} (RMET)_{ij} + [r_{0j} + e_{ij}],$$

such that  $HS_{ij}$  was the beginning-of-semester helping skills variable,  $\pi_{00}$  was the overall intercept, which was the average of the beginning-of-semester helping skills, and  $\pi_{10}$ ,  $\pi_{20}$ ,  $\pi_{30}$ ,  $\pi_{40}$ , and  $\pi_{50}$  were the main effects for the five

student-level variables (IRI, NHM, PHE, FIS, and RMET, respectively). The parameters inside the  $[\ ]$  bracket were the random effects. In this model,  $r_{0j}$  accounted for the instructor variability around the overall intercept ( $\pi_{00}$ ).

Prediction of late-semester helping skills outcomes was similar, except the beginning-of-semester level of helping skills was entered as a covariate. The five predictors were the same as in the first analysis. Specifically, the model was

$$\begin{aligned} \text{HSTime2}_{ij} = & \pi_{00} + \pi_{10} (\text{HSTime1})_{ij} + \pi_{20} (\text{IRI})_{ij} + \pi_{30} (\text{NHM})_{ij} \\ & + \pi_{40} (\text{PH})_{ij} + \pi_{50} (\text{FIS})_{ij} + \pi_{60} (\text{RMET})_{ij} + [r_{0j} + e_{ij}], \end{aligned}$$

such that  $\text{HSTime2}_{ij}$  was the late-semester helping skills variable,  $\text{HSTime1}_{ij}$  was the beginning-of-semester helping skills variable which served as a covariate,  $\pi_{00}$  was the overall intercept which was the average of the beginning-of-semester helping skills, and  $\pi_{20}$ ,  $\pi_{30}$ ,  $\pi_{40}$ ,  $\pi_{50}$ , and  $\pi_{60}$  were the main effects for the five student-level variables (IRI, NHM, PHE, FIS, and RMET, respectively). The parameters inside the bracket were the random effects. Similar to the previous model,  $r_{0j}$  accounted for the instructor variability around the overall intercept ( $\pi_{00}$ ). Thus, the model tested whether the five independent variables predicted the late-semester levels of helping skills after controlling for the beginning-of-semester levels of helping skills.

## Results

For questionnaires in which less than 10% of the items were missing values, the missing values were replaced using linear interpolation. No problems were found when we checked for normality and outliers using observations of histograms and assessments of skewness and kurtosis.

No differences were found between the participants in the fall and spring semesters in the measures administered at the beginning of the semester, suggesting that the different recruiting methods and participation rate did not influence the data. However, there were instructor differences for SPOM-Hr at both beginning-of-semester,  $F(3, 160) = 3.84, p = .01$ , and late-semester,  $F(3, 146) = 3.23, p = .02$ . There were also instructor differences for the proportion of words spoken at the beginning-of-semester,  $F(3, 182) = 7.30, p < .001$ , and late-semester,  $F(3, 181) = 12.07, p < .001$ . These instructor differences were found to be from one instructor whose students, relative to all other students, had lower SPOM-Hr scores at both the beginning-of-semester,  $F(1, 162) = 10.33, p = .002$ , and late-semester,  $F(1, 148) = 7.75, p = .006$ , and a lower proportion of words spoken at the beginning-of-semester,  $F(1, 183) = 15.07, p < .001$ , and late-semester,  $F(1, 184) = 20.67, p < .001$ .

For variables where there were instructor differences, an instructor variable was created, which coded students taught by this single instructor as one level and all other students as the second level. The bifurcated instructor variable was included as a factor in the relevant tests of helping skills changes from beginning-of-semester and late-semester. For the second set of analyses, involving the prediction of helping skills, we accounted for the intraclass variation within instructors by assigning instructor as a random effect within a hierarchical analysis.

### *Effects of Helping Skills Training*

Table 1 shows the means and standard deviations for all the measures used in this study. Table 2 shows correlations among all measures for the present study. Using effect size analyses (differences between means divided by pooled standard deviations), this sample was not significantly different from other samples (Chui et al., 2014; Hill & Kellems, 2002; Hill et al., 2008; Jackson et al., 2014; Spangler et al., 2014) on the NHM, PHE, CASES-HS, SPOM-C, SPOM-Hr, exploration skills, and helper words.

One set of repeated measures ANOVA was conducted for SPOM-C, proportion of exploration skills, and CASES-HS. Significant changes occurred across time for all three measures (see Table 1). For SPOM-Hr and helper words (which both had instructor differences in the preliminary analysis), we conducted a  $2 \times 2$  mixed measures ANOVAs with time as the repeated factor and instructor as the between-subjects factor. In these analyses, significant main effects were found for time, but no significant instructor or Instructor  $\times$  Time effects were found. Hence, it appeared that students changed over the course of the semester on all five outcome measures, and this change was not influenced by instructor. As seen in Table 1, effect sizes varied, such that the effect for CASES was large; the effect for helper words spoken was moderate; and the effects for SPOM-Hr, SPOM-C, and exploration interventions were small.

### *Prediction of Beginning- and Late-Semester Helping Skills*

Predictors of helping skills were analyzed with two sets of HLM analyses (Raudenbush & Bryk, 2002), where students were nested within instructors. As noted previously, the predictors of IRI, NHM, PHE, FIS, and RMET were entered simultaneously to test each of the five helping skill variables for (a) beginning-of-semester levels of helping skills and (b) late-semester helping skills. For both sets of HLM analyses, the nested effects for instructor did not emerge as significant for either model.



**Table 1.** Means, Standard Deviations, and Training Effects for Helping Skills Variables.

	Beginning-of-semester			Late-semester			F	$\eta^2$
	n	M	SD	n	M	SD		
<b>Predictors</b>								
IRI	191	58.21	8.71					
NHM	191	5.53	1.02					
PHE	191	1.96	0.90					
FIS	164	2.93	0.46					
RMET	191	26.80	3.00					
<b>Helping skills</b>								
CASES-HS	125	4.47	1.34	125	6.57	1.07	381.03***	.75
SPOM-C	149	4.06	0.57	149	4.28	0.40	19.99***	.12
SPOM-Hr	146	3.88	0.51	146	4.06	0.44	19.58***	.12
Explore	180	0.70	0.16	180	0.76	0.12	18.38***	.09
Words	179	0.32	0.12	179	0.22	0.09	123.81***	.41

Note.  $\eta^2$  refers to the effect size estimate partial eta-squared. IRI = Interpersonal Reactivity Index; NHM = Natural Helping Measure; PHE = Prior Helping Experiences Measure; FIS = Facilitative Interpersonal Skills; RMET = Reading of Mind in Eyes Test; CASES-HS = Counseling Activity Self-Efficacy Scales–Helping Skills; SPOM-C = Session Process and Outcome Measure–Client; SPOM-Hr = Session Process and Outcome Measure–Helper; Explore = exploratory skills; Words = helper words (1 and 2 refers to session numbers). \*\*\* $p < .001$ .

**Prediction of beginning-of-semester helping abilities.** The IRI emerged as the sole predictor for four of the five helping abilities: SPOM-C,  $\beta = .02$  ( $SE = .005$ ),  $t(127) = 3.68$ ,  $p < .001$ ; SPOM-Hr,  $\beta = .01$  ( $SE = .005$ ),  $t(128) = 2.10$ ,  $p = .047$ ; CASES,  $\beta = .04$  ( $SE = .014$ ),  $t(100) = 2.53$ ,  $p = .013$ ; and helper words spoken,  $\beta = -.003$  ( $SE = .001$ ),  $t(151) = -2.88$ ,  $p = .004$ . There were no significant predictors for initial level of exploratory interventions. Thus, higher IRI scores predicted higher client and helper ratings of session quality, higher levels of helper-rated self-efficacy, and lower overall helper words spoken. These results suggest that empathy (as measured by the IRI) significantly predicted levels of helping skills at the beginning of training. In other words, a student's entry-level of trait empathy appeared to be the best indicator of pretraining levels of most of the helping skills.

**Prediction of late-semester helping abilities.** Most predictive of late-semester helping abilities were the beginning-of-semester levels of the same helping abilities. Thus, all of the five late-semester helping skills were significantly

**Table 2.** Correlations of Predictor and Helping Skill Criterion Variables.

	IRI	NHM	PHE	RMET	FIS	CASES-P	CASES-N	SPOM-C1	SPOM-C2	SPOM-Hr1	SPOM-Hr2	Explore 1	Explore 2	Words 1	Words 2
IRI	—														
NHM	.36***	—													
PHE	.13*	.33***	—												
RMET	.04	.01	.06	—											
FIS	.07	.19**	.08	.05	—										
CASES-P	.28***	.22**	.24**	.02	.13	—									
CASES-N	.20**	.30***	.15*	.03	.19**	.52***	—								
SPOM-C1	.36***	.22**	.24**	.02	.13	.31***	.18*	—							
SPOM-C2	.02	.02	.05	.01	-.01	.15*	.22**	.26***	—						
SPOM-Hr1	.21**	.26***	.19*	-.11	.03	.25**	.28***	.39***	.20**	—					
SPOM-Hr2	.17*	.28***	.16*	-.05	.09	.28***	.33**	.38***	.52***	.38***	—				
Explore 1	-.01	-.01	.10	.17*	.12	-.16*	.03	-.04	-.04	-.22**	-.06	—			
Explore 2	.10	.01	-.02	.08	-.02	-.08	.00	.15*	-.03	-.05	-.03	.19**	—		
Words 1	-.22**	-.03	.09	-.06	.01	.14	-.01	.09	.13	.23**	.01	-.38***	-.05	—	
Words 2	-.11	.06	.14*	.02	.04	.12	-.09	.03	.16*	.14*	.13	-.11	-.09	.37***	—

Note. IRI = Interpersonal Reactivity Index; NHM = Natural Helping Measure; PHE = Prior Helping Experiences Measure; RMET = Reading of Mind in Eyes Test; FIS = Facilitative Interpersonal Skills; CASES-P = Counseling Activity Self-Efficacy Scales-Pretraining; CASES-N = Counseling Activity Self-Efficacy Scales-Now; SPOM-C = Session Process and Outcome Measure-Client; SPOM-Hr = Session Process and Outcome Measure-Helper; Explore = exploratory skills; Words = helper words (1 and 2 refers to session numbers). \* $p < .05$ . \*\* $p < .01$ . \*\*\* $p < .001$ .

predicted by the pretraining levels of these skills (all in the predicted direction): SPOM-C,  $\beta = .20$  ( $SE = .066$ ),  $t(128) = 3.05$ ,  $p = .003$ ; SPOM-Hr,  $\beta = .40$  ( $SE = .068$ ),  $t(125) = 5.79$ ,  $p < .001$ ; CASES,  $\beta = .38$  ( $SE = .072$ ),  $t(101) = 5.28$ ,  $p < .001$ ; exploratory interventions,  $\beta = .13$  ( $SE = .059$ ),  $t(154) = 2.23$ ,  $p = .027$ ; and helper words spoken,  $\beta = .20$  ( $SE = .053$ ),  $t(150) = 3.71$ ,  $p < .001$ .

From the full multilevel model, FIS predicted CASES (after covarying out beginning-of-semester CASES),  $\beta = .40$  ( $SE = .188$ ),  $t(101) = 2.12$ ,  $p = .037$ . None of the other independent variables (IRI, NHM, PHE, or RMET) were significant predictors of late-semester helping skill outcomes. Thus, end-of-semester performance was related with beginning-of-semester performance, such that those who started high ended high and those who started low ended low; performance on FIS (the simulated helping task) was related to end levels of self-efficacy.

## Discussion

Replicating previous results (Hill & Kellems, 2002; Hill et al., 2008), we demonstrated the effectiveness of Hill's (2009, 2014) model of helping skills training for undergraduate students. More specifically, undergraduate students improved in terms of their helping abilities, which we defined as increased self-efficacy for using the helping skills, reduced talking in sessions, implementing more exploration skills in sessions, and conducting better quality sessions (as judged by themselves and volunteer clients). Thus, we can state with some confidence that undergraduates are capable of learning Hill's helping skills model, although we do not know whether these results would extend to training involving other models. We also note that after one semester of training, these undergraduate students were not yet able to function in a counseling situation. They were still at a novice level in terms of their ability to implement the skills in highly structured situations and had yet to gain knowledge and skills related to other therapeutic abilities (e.g., therapeutic presence, self-awareness, case conceptualization, knowledge of ethics, multicultural awareness, theory). Given the support for the effectiveness of helping skills training, we then examined whether we could predict helping abilities at the beginning and end of training.

### *Predicting Helping Abilities Prior to Training*

Several measures of the study, chosen because they appeared to characterize interpersonal ability, were significantly correlated with each other at the beginning of the semester. Thus, when students thought they were empathic

and natural helpers, they also reported having sought out more prior helping experiences, and they had high self-efficacy for using the helping skills (based on retrospective pre-ratings).

Interestingly, the only predictor variable that was *uniquely* related to the performance measures of helping ability prior to training was the self-report measure of empathy (the IRI). Thus, when trainees viewed themselves as empathic, they performed better in brief helping sessions (i.e., were judged by themselves and their volunteer clients as conducting high-quality sessions, did not speak a lot in sessions, and had high retrospective prelevels of self-efficacy for using the helping skills). Perhaps when students thought of themselves as empathic, they were more relaxed and able to be present with the client in the brief helping session. These results suggest that trainees are relatively aware of their empathic abilities.

Surprisingly, the FIS task, a simulated performance-based assessment, was not related prior to training to any of the other measures of interpersonal ability or to the actual helping behaviors in sessions (exploration skills, words spoken), to client-rated or helper-rated evaluations of sessions (SPOM-C and SPOM-Hr), nor to retrospective preassessments of self-efficacy for using helping skills. Apparently, the ability to respond spontaneously to a video clip of a difficult client is different from lay helpers' self-assessments and ability to be helpful in a brief session. Students may have felt more motivated to be helpful with a peer serving as a client than they did in responding to the video clips in the FIS task. It also may have been difficult for untrained students to respond to the unusual and difficult therapy scenarios in the video clips, whereas they were used to responding empathically to peers.

It was also surprising that the amount of exploration skills used in the helping session was not related to other beginning-of-semester measures of helping ability. Perhaps using exploration skills is not indicative of one's helping ability. These results support the notion that one can use exploration skills but not be empathic, as suggested by Hill (2014).

### *Prediction of Helping Abilities at the End of Training*

The strongest predictor of students' helping abilities at the end of training was presemester levels of helping abilities, such that past behavior predicted future behavior. Thus, although the average student gained in helping skills, those who started as the best helpers were still the best helpers at the end of the semester, and those who started as the worst helpers were still the worst helpers at the end of the semester. In other words, if one wanted to predict who would benefit most from training, client- and helper-rated SPOM scores from an initial session would be a relatively good performance-based measure for

doing so. However, the magnitude of beginning-to-late predictions correlations were moderate, meaning that there was room for improvement for all students as a result of training. Taken together, the results not only point to the advantages of being skilled before training but also suggest that trainees can change through training and practice.

The only other predictor of end of semester outcomes was that performance on the FIS task predicted self-efficacy after controlling for retrospective preassessment of self-efficacy and instructor effects. Thus, those students who were judged as responding effectively to video clips of difficult clients at the beginning of the semester were most confident of their ability to use helping skills after training (although as noted previously, performance on the FIS task was not related to retrospective prelevels of self-efficacy). Perhaps responding effectively to the video clips (FIS) requires a certain level of confidence, an ability to think quickly and respond assertively, which is similar to what is required in learning helping skills.

### *Limitations and Strengths*

The participants were all from one public university in the mid-Atlantic United States, which limits generalizability to other settings. Furthermore, participants self-selected into taking the helping skills course; they were psychology majors who had to take a laboratory course, although there were other labs that they could have chosen. Although these participants were similar to those from previous studies at the same university on outcome measures, the current participants may not represent the entire range of scores on the various measures (e.g., we might expect that they were more empathic than the average non-psychology major).

Another limitation relates to the instructors. All were at the same university and had been trained by the same professor in helping skills. In addition, all were aware of the general purposes of the study. Interestingly, there was an instructor effect on two of the outcome measures. One instructor had lower scores on both the SPOM-Hr and proportion of words, but these were at both the beginning and end of the semester, suggesting that these effects may have been due to another class-related factor (e.g., students, time of day) rather than the instructor. Moreover, the multilevel modeling with students nested within instructors did not find instructor effects in any of the statistical models, suggesting that instructor effects did not substantially influence the results. Although we had an adequate sample size of helpers, there were only six classes, and students were nested within classes. Due to the small number of classes and instructors, we could not test for what accounted for differences among instructors and classes.

An additional limitation involves the use of classmates as volunteer clients. Although using classmates is easier when dealing with large numbers of students as in the present study and because it is easier to monitor any potential harmful effects of a student who is a very ineffective helper, there are some drawbacks. Classmates learn the skills too and thus are not naïve at the end of the semester. Also, a better representation of one's ability would be to have several clients at both the beginning and end, to control for individual differences among clients.

A strength of the study was that we used multiple methods (self-report, nonverbal measures, performance measures) for assessing the predictors and outcomes. There was not enough consistency in these measures to remove method variance, but method variance should be taken into account for future large-scale studies. Finally, the FIS ratings differed from previous studies in that there was a higher internal consistency in the ratings of the eight items. Hence, raters seemed to be rating a core FIS construct rather than differentiating eight constructs as judges have done in previous studies.

### *Implications*

The replication of findings for the effectiveness of helping skills training using the Hill model across several studies (Chui et al., 2014; Hill & Kellems, 2002; Hill et al., 2008; Jackson et al., 2014; Spangler et al., 2014) provides support for this model as an evidence-based training model for training undergraduate students in helping skills. Further research on teaching exploration and action stage skills, instructor effects, and use of technology would enhance the model even more.

Given the paucity of studies predicting outcomes of helping skills training, it is premature to state definitive implications for selection of therapists from these data. At the present time, however, if called upon to select potential counselors from an undergraduate population, we would suggest selecting those scoring high on self-reports of empathy, scoring high on self-ratings and volunteer client ratings of session quality from a brief helping session, having a low proportion of words spoken in a brief helping session, and having high self-report estimates of self-efficacy for using the helping skills. Nonetheless, it appears that most students can benefit from training, so there is no empirical reason to exclude anyone from training.

Future researchers could attempt to replicate the finding of empathy as a predictor of helping abilities prior to training. We recommend that future researchers use large samples as this study's larger sample (by comparison with past studies) may have made it possible for this relatively weak effect to

become more robust. More research is also needed to understand how self-efficacy, interpersonal performance, and demonstrated helping skills are related at various points in the training process.

We also suggest that future researchers refine the use of the brief helping sessions as a prediction tool. Perhaps students could conduct sessions with three or more volunteer clients to obtain a more stable estimate of their abilities. In addition, future researchers could use volunteer clients other than classmates as a better assessment of ability to use skills, though we recognize the thorny ethical issues that this practice may raise, including the need for close supervision to prevent harm to volunteer clients. In addition, perhaps a version of the FIS task could be created that would be more applicable to undergraduate students and the types of clients with whom they might be working. Finally, it would be useful to include additional measures of training outcome, including perhaps some of the predictor measures that we used at the beginning of the semester to determine whether trainees change in terms of interpersonal functioning.

### Declaration of Conflicting Interests

The authors declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

### Funding

The authors received no financial support for the research, authorship, and/or publication of this article.

### References

- Anderson, T., Crowley, M. J., Himawan, L., Holmberg, J., & Uhlin, B. (2015). Therapist facilitative interpersonal skills and training status: A randomized clinical trial on alliance and outcome. *Psychotherapy Research*. Advance online publication. doi:10.1080/10503307.2015.1049671
- Anderson, T., McClintock, A. S., Himawan, L., Song, X., & Patterson, C. L. (in press). A prospective study of therapist facilitative interpersonal skills as a predictor of treatment outcome. *Journal of Consulting and Clinical Psychology*.
- Anderson, T., Ogles, B. M., Patterson, C. L., Lambert, M. J., & Vermeersch, D. A. (2009). Therapist effects: Facilitative interpersonal skills as a predictor of therapist success. *Journal of Clinical Psychology, 65*, 755-768. doi:10.1002/jclp.20583
- Austin, E. J. (2004). An investigation of the relationship between trait emotional intelligence and emotional task performance. *Personality and Individual Differences, 36*, 1855-1864. doi:10.1016/j.paid.2003.07.006
- Baker, S. B., & Daniels, T. G. (1989). Integrating research on the microcounseling program: A meta-analysis. *Journal of Counseling Psychology, 36*, 213-222. doi:10.1037/0022-0167.36.2.213

- 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. doi:10.1177/0011000090183001
- Baron-Cohen, S., Wheelwright, S., Hill, J., Raste, Y., & Plumb, I. (2001). The "reading the mind in the eyes" test revised version: A study with normal adults, and adults with Asperger syndrome or high-functioning autism. *Journal of Child Psychology and Psychiatry and Allied Disciplines, 42*, 241-251. doi:10.1111/1469-7610.00715
- 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.
- Chemers, M. M., Hu, L., & Garcia, B. F. (2001). Academic self-efficacy and first-year college student performance and adjustment. *Journal of Educational Psychology, 93*, 55-64. doi:10.1037/0022-0663.93.1.55
- Chui, H., Hill, C. E., Ain, S., Ericson, S., Del Pino, H. G., Hummel, A., . . . Spangler, P. T. (2014). Training undergraduate students to use challenges. *The Counseling Psychologist, 42*, 758-777. doi:10.1177/0011000014542835
- Cook, C. M., & Saucier, D. M. (2010). Mental rotation, targeting ability and Baron-Cohen's empathizing systemizing theory of sex differences. *Personality and Individual Differences, 49*, 712-716. doi:10.1016/j.paid.2010.06.010
- Davis, M. H. (1980). 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.
- Dunning, D., Heath, C., & Suls, J. M. (2004). Flawed self-assessment implications for health, education, and the workplace. *Psychological Science in the Public Interest, 5*, 69-106. doi:10.1111/j.1529-1006.2004.00018.x
- 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. doi:10.1037/0022-006X.55.2.218
- Ford, J. D. (1979). Research on training therapists and clinicians. *Review of Educational Research, 49*, 87-130.
- Gallese, V., Keysers, C., & Rizzolatti, G. (2004). A unifying view of the basis of social cognition. *Trends in Cognitive Sciences, 8*, 396-403. doi:10.1016/j.tics.2004.07.002
- 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. doi:10.1080/10503300500264838
- Hill, C. E. (1978). Development of a counselor verbal response category system. *Journal of Counseling Psychology, 25*, 461-468. doi:10.1037/0022-0167.25.5.461
- 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, NY: Guilford.



- Hill, C. E. (2004). *Helping skills: Facilitating exploration, insight, and action* (2nd ed.). Washington, DC: American Psychological Association.
- Hill, C. E. (2009). *Helping skills: Facilitating exploration, insight, and action* (3rd ed.). Washington, DC: American Psychological Association.
- Hill, C. E. (2014). *Helping skills: Facilitating exploration, insight, and action* (4th 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. doi:10.1037/0022-0167.49.2.264
- Hill, C. E., & Knox, S. (2013). Training and supervision in psychotherapy: Evidence for effective practice. In M. J. Lambert (Ed.), *Handbook of psychotherapy and behavior change* (6th ed., pp. 775-811). New York, NY: 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. doi:10.1037/0033-3204.43.2.154
- Hill, C. E., & O'Brien, K. (1999). *Helping skills: Facilitating exploration, insight, and action*. Washington DC: American Psychological Association.
- Hill, C. E., Roffman, M., Stahl, J., Friedman, S., Hummel, A., & Wallace, C. (2008). Helping skills training for undergraduates: Outcomes and predictors of outcomes. *Journal of Counseling Psychology, 55*, 359-370. doi:10.1037/0022-0167.55.3.359
- Hill, C. E., Spangler, P. T., Jackson, J., & Chui, H. (2014). Training undergraduate students to use insight skills: Integrating results across three studies. *The Counseling Psychologist, 42*, 800-820. doi:10.1177/0011000014542602
- 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.
- Jackson, J., Hill, C. E., Spangler, P. T., Ericson, S., Merson, E., Liu, J., . . . Reen, G. (2014). Training undergraduate students to use interpretation. *The Counseling Psychologist, 42*, 778-799. doi:10.1177/0011000014542600
- 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, L. (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: Thomas.
- Lambert, M. J., DeJulio, S. S., & Stein, D. M. (1978). Therapist interpersonal skills: Process, outcome, methodological considerations, and recommendations for future research. *Psychological Bulletin, 85*, 467-489.
- Lawrence, E. J., Shaw, P. P., Baker, D. D., Baron-Cohen, S. S., & David, A. S. (2004). Measuring empathy: Reliability and validity of the empathy quotient. *Psychological Medicine, 34*, 911-919. doi:10.1017/S0033291703001624

- 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. doi:10.1037/0022-0167.50.1.97
- Mabe, P. A., III, & West, S. G. (1982). Validity of self-evaluation of ability: A review and meta-analysis. *Journal of Applied Psychology, 67*, 280-286. doi:10.1037//0021-9010.86.4.730
- Matarazzo, R. G. (1971). The systematic study of learning psychotherapy skills. In A. E. Bergin & S. L. Garfield (Eds.), *Handbook of psychotherapy and behavior change* (pp. 895-924). New York: Wiley.
- Matarazzo, R. G. (1978). Research on the teaching and learning of psychotherapeutic skills. In S. L. Garfield & A. E. Bergin (Eds.), *Handbook of psychotherapy and behavior change* (2nd ed., pp. 941-966). New York: Wiley.
- Matarazzo, R. G., & Patterson, D. (1986). Research on the teaching and learning of therapeutic skills. In S. L. Garfield & A. E. Bergin (Eds.), *Handbook of psychotherapy and behavior change* (3rd ed., pp. 821-843). New York: Wiley.
- McDaniel, M. A., Morgeson, F. P., Finnegan, E. B., Campion, M. A., & Braverman, E. P. (2001). Use of situational judgment tests to predict job performance: A clarification of the literature. *Journal of Applied Psychology, 86*, 730-740. doi:10.1037/0021-9010.86.4.730
- Ogles, B. M. (2013). Measuring change in psychotherapy research. In M. J. Lambert (Ed.), *Handbook of psychotherapy and behavior change* (6th ed., pp. 134-166). New York, NY: Wiley.
- Padykula, N., & Horwitz, M. (2012). Using psychodynamic concepts to measure interpersonal competencies during social work training. *International Journal of Applied Psychoanalytic Studies, 9*, 48-61. doi:10.1002/aps.312
- Raudenbush, S., & Bryk, A. (2002). *Hierarchical linear models* (2nd ed.). Newbury Park, CA: SAGE.
- Ridley, C. R., Kelly, S. M., & Mollen, D. (2011). Microskills training: Evolution, reexamination, and call for reform. *The Counseling Psychologist, 39*, 800-824. doi:10.1177/0011000010378438
- Rogers, C. R. (1957). The necessary and sufficient conditions of therapeutic personality change. *Journal of Consulting Psychology, 21*, 95-103. doi:10.1037/h0045357
- Russell, R. R., Crimmings, A. M., & Lent, R. W. (1984). Therapist training and supervision. In S. Brown & R. W. Lent (Eds.), *Handbook of Counseling Psychology* (pp. 625-681). New York: Wiley.
- Singer, J. (1998). Using SAS PROC MIXED to fit multilevel models, hierarchical models, and individual growth models. *Journal of Educational and Behavioral Statistics, 24*, 323-355. doi:10.3102/10769986023004323
- Spangler, P. T., Hill, C. E., Dunn, M. G., Hummel, A., Walden, T., Liu, J., . . . Salahuddin, N. (2014). Training undergraduate students to use immediacy. *The Counseling Psychologist, 42*, 729-757. doi:10.1177/0011000014542835
- Stahl, J. V., & Hill, C. E. (2008). A comparison of four methods for assessing natural helping ability. *Journal of Community Psychology, 36*, 289-298. doi:10.1002/jcop.20195

- Stajkovic, A. D., & Luchins, F. (1998). Self-efficacy and work-related performance: A meta-analysis. *Psychological Bulletin, 124*, 240-261. doi:10.1037/0033-2909.124.2.240
- Strupp, H. H. (1993). The Vanderbilt psychotherapy studies: Synopsis. *Journal of Consulting and Clinical Psychology, 61*, 431-433.

## Author Biographies

**Clara E. Hill**, PhD, is a professor of psychology in the counseling psychology program at the University of Maryland. Her research areas of interest are the identification and training of helping skills, process and outcome of psychotherapy, working with dreams, qualitative research, and meaning in life.

**Timothy Anderson** is an associate professor of psychology in the clinical psychology program at Ohio University. His research focuses on the common elements of psychotherapy, including therapist interpersonal characteristics, expectations, and relational processes.

**Kathryn Kline** is a doctoral student in the counseling psychology program at the University of Maryland. Her research interests are the process and outcome of psychotherapy, therapist training, and meaning in life.

**Andrew McClintock** is a doctoral student in the clinical psychology program at Ohio University. His research program focuses on the process and outcome of psychological interventions, particularly cognitive- and mindfulness-based treatments.

**Saryn Cranston** is a doctoral student in the clinical psychology program at Ohio University. Her research areas of interest are corrective experiences and change in psychotherapy.

**Shannon McCarrick** is a doctoral student in clinical psychology at Ohio University. Her research interests include psychotherapy process and outcome with a focus on common factors, psychotherapy with LGBT persons, and personality disorders and attachment.

**Allison Petrarca** is a doctoral student in the clinical psychology program at Ohio University. Her research interests include psychotherapy outcomes, mindfulness, and how religious beliefs impact obsessive-compulsive disorder.

**Lina Himawan** is a biostatistician in the department of psychology at Ohio University. She assists faculty members and graduate students in the department with writing grant proposals, conducting power analyses, selecting research designs, conducting data analyses, and presenting results of analyses for publication.

**Andrés E. Pérez-Rojas**, PhD, is a postdoctoral fellow at the Center for Counseling and Psychological Services at the Pennsylvania State University. He received his doctoral degree in counseling psychology from the University of Maryland, College Park. His main research interests focus on multiculturalism, particularly on topics

related to bilingual Latinos and international students, and psychotherapy process and outcome, with an emphasis on the therapeutic relationship, therapist effects, and therapist training.

**Avantika Bhatia** is a doctoral student in the counseling psychology program at the University of Maryland. Her research areas of interest are process and outcome of psychotherapy, including the therapeutic relationship and treatment termination.

**Shudarshana Gupta** is a doctoral student in the counseling psychology program at the University of Maryland. Her research areas of interest are the identification and training of helping skills, the role of laughter in psychotherapy, working with dreams, and process and outcome of psychotherapy.

**Margo Gregor** is a doctoral student in the counseling psychology program at the University of Maryland. Her research interests focus on the vocational development of women and issues surrounding training and supervision of clinicians.