The Counseling Psychologist
2014, Vol. 42(6) 702–728

DOI: 10.1177/0011000014542598
tcp.sagepub.com
SAGE

Training Undergraduate Students to Use Insight Skills: Overview of the Rationale, Methods, and Analyses for Three Studies

Clara E. Hill[1], Patricia T. Spangler[1], Harold Chui[1], and John L. Jackson[1]

Abstract

We briefly review the literature on helping skills training. We then provide a rationale for the current series of studies, given methodological problems and a lack of focus on teaching insight skills in the previous literature. Next, we provide an overview of the rationale, methods, and analyses used in common across three studies conducted to teach insight skills (immediacy, challenges, and interpretation, respectively) to undergraduate students in helping skills courses.

Keywords

helping skills, insight, counselor training, mixed-method, self-efficacy

For many students wanting to become therapists, their first formal training experience involves learning helping skills in an undergraduate helping skills course. Given the foundational nature of this training, it is important to have evidence about the effectiveness of the entire training program as well as the effectiveness of the components of training.

[1]University of Maryland, College Park, MD, USA

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

As reviewed next, we have considerable evidence (albeit methodologically limited) for the effectiveness of training programs involving basic or exploration skills, but we have minimal evidence about the effectiveness of teaching advanced or insight skills. Given that insight skills are difficult to teach because they require some mastery of exploration skills and theoretical knowledge of the target skills, we need empirical research so that we can improve our training methods.

In this series of three studies, we tested the effectiveness of training for three different skills (immediacy, challenge, and interpretation) taught in the insight stage of the Hill (2009) helping skills model. To set the stage for these three studies, we review the history of helping skills training research as well as the more recent research on the Hill training model. We then present the methodological issues, methods, and analyses that were common across the three studies.

History of Helping Skills Training in Counseling Psychology

Helping skills training has been a part of counseling psychology since the beginning of the profession (Hill & Knox, 2013; Hill & Lent, 2006). Based on Rogers's (1942) early theorizing about the importance of serving as a sounding board or mirror so that clients could hear themselves and explore more deeply, many programs were developed for teaching helping skills to professionals, paraprofessionals, teachers, and parents.

Effectiveness of Early Training Programs

The early programs that received the most empirical attention were Carkhuff's (1969) Human Relations Training (HRT), Ivey's (1971) Microcounseling (MC), and Kagan's (1984) Interpersonal Process Recall (IPR). In HRT and MC, trainees are taught specific helping skills (e.g., restatement, reflection of feeling) in a structured group format. In IPR, in contrast, trainees are helped to articulate their thoughts and feelings about their interventions under the assumption that trainees can effectively use their natural communication skills if they can become aware of the blocks from performance anxiety. A number of narrative reviewers concluded that helping skills training was effective (Ford, 1979; Hill & Knox, 2013; Kasdorf & Gustafson, 1978; Lambert, DeJulio, & Stein, 1978; Matarazzo, 1971, 1978; Matarazzo & Patterson, 1986; Russell, Crimmings, & Lent, 1984). Matarazzo, for example, suggested that there was sufficient evidence for the effectiveness of teaching Ivey's MC program.

In a meta-analysis, Baker and Daniels (1989) found a large effect size for undergraduate trainees (1.18) and a moderate effect size for graduate trainees (0.66) for Ivey's MC training. In an additional meta-analysis comparing MC, HRT, and IPR for graduate trainees, Baker, Daniels, and Greeley (1990) found a large effect (1.07) for Carkhuff's HRT, a medium effect (0.63) for Ivey's MC, and a small effect (0.20) for Kagan's IPR.

A number of the reviews (e.g., Hill & Knox, 2013; Hill & Lent, 2006; Ridley, Mollen, & Kelly, 2011), however, described methodological problems with the early research on helping skills training that limit our confidence in the findings: (a) definition and assessment of the skills was poor; (b) details of the training were rarely specified; (c) researchers assessed the effects of training using written or spoken responses to written or audiotaped analog clients, a method that does not produce the same results as responding to a client in an actual helping setting; (d) the same rating scales were used to assess training and outcome, and it often appeared that trainers were just teaching to the test; (e) because only one trainer was used in most of the studies, researchers could not separate out the effects of the training from the specific trainer; (f) most studies used very small samples of trainees and did not control for group effects; (g) control conditions were not typically employed; and (h) the relation of skill use to client outcome has not typically been investigated. In addition, the skills that typically have been taught in past studies were empathy or reflection of feelings (exploration skills), which were often taught outside the framework of training (e.g., brief 30-min training of just one skill). These skills are relatively easy to teach in a brief structured format if the goal of training is simply to get the trainee to formulate a grammatically correct approximation of the skill (although these skills are not easy to apply clinically). Notably, complex skills such as immediacy, challenge, and interpretation are difficult to teach in a brief format outside the context of a training program because these advanced skills are not typically taught until later in training after students have learned exploration skills.

Effectiveness of the Hill Model of Helping Skills Training

A majority of more recent helping skills research has focused on the Hill helping skills model (Hill, 2004, 2009, 2014; Hill & O'Brien, 1999), which Ridley et al. (2011) evaluated 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 relation of skills to therapeutic change. This 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 has been 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.

In this three-stage model, there are goals and skills for each stage. The goals of the exploration stage are to build the relationship, help clients tell their stories, and help clients explore thoughts (through open questions for thoughts and restatements) and feelings (through open questions for feelings, reflections of feelings, and disclosures of feelings). The goals of the insight stage are to facilitate awareness (through challenges), promote insight (through open questions for insight, interpretations, and disclosure of insights), and facilitate insight into the therapeutic relationship (through immediacy). The goals of the action stage are to help the client in the change process, using the skills of open questions for action, disclosure of strategies, information, and direct guidance combined into working on four different types of changes (relaxation, behavior change, behavior rehearsal, and decision making). Importantly, the skills are not taught to be applied in a rigid manner, but rather students are taught that a variety of skills can be helpful for different goals, and that the manner in which skills are implemented is crucial. Thus, a major emphasis is made to teach empathy, collaboration, flexibility, clinical intuition, ethics, self-awareness, and awareness of client reactions along with practicing the individual skills.

In undergraduate training, students are generally first required to read the relevant chapter of the helping skills text (e.g., Hill, 2009, 2014). In class, instructors then provide a theoretical rationale for each skill within each stage and facilitate a discussion to stimulate thinking, so that students can learn to challenge assumptions and think for themselves. Instructors then provide examples through videos or demonstrations to model for the trainees how the skill can be used. Finally, instructors provide students with a variety of opportunities to practice the skills, often in small groups or dyads, offering support, modeling, coaching, and feedback during the practice.

Given that a major criticism of the early helping skills studies was that training outcome was assessed by trainees only responding to written or audio analog vignettes, it seemed important to develop measures that could assess trainee performance in more realistic clinical settings. Hill and Kellems (2002) thus developed the Helping Skills Measure (assesses the extent to which exploration, insight, and action skills were perceived as being used), the Relationship Scale (assesses the perceived quality of the therapeutic relationship), and the Session Evaluation Scale (assesses the perceived quality of the session). Helpers were undergraduate and graduate students in helping skills classes. The undergraduate helpers were initially paired with a classmate for a 20-min helping session and then had a 20-min helping session with an

introductory psychology volunteer "client" in mid-semester and a 45-min helping session with a different "client" at the end of the semester. Graduate student therapists conducted two to five helping sessions with introductory psychology students who served as volunteer "clients." Helpers and "clients" completed the three measures after sessions. Hill and Kellems found that both undergraduate and graduate student helpers improved over the course of a semester. More specifically, volunteer clients at the end of the semester perceived the trainees as using more exploration, insight, and action skills than at the beginning of the semester. In addition, trainees were rated as providing a better therapeutic relationship, and sessions were rated as higher in quality.

Hill et al. (2008) expanded on the Hill and Kellems (2002) study by investigating the effects of helping skills training from the perspectives not only of helpers and clients but also of judgments of behavior in sessions. Based on brief sessions with classmates before and after training in the exploration stage, Hill et al. found that undergraduate trainees used more exploration skills (as assessed by perceptions of clients and helpers as well as by counts of behaviors in sessions), were judged as being more empathic, talked less of the time in the session, and were judged as more effective after training in exploration skills than they were before the training. Moreover, in post-semester self-evaluations, trainees indicated that they had higher self-efficacy for using helping skills than they had prior to training. In terms of confidence for using the skills as rated after the weekly lab, trainees indicated that their confidence increased steadily during training in exploration skills, decreased somewhat during training in insight skills, and increased to the highest level after training in action skills. Finally, initial grade point average, self-rated empathic concern and perspective taking ability, and self-rated perfectionism did not predict who benefited from training. Thus, these results showed fairly convincing evidence that undergraduate students did indeed learn to use helping skills, although the authors were not able to predict who benefited from training.

The results from these studies with undergraduate trainees suggest that training in the Hill helping skills model is effective. These preliminary results are encouraging for showing the overall effectiveness of the training model, but we need to know more about training for specific skills, the effectiveness of components of the model, and whether we can predict who benefits from training.

The Effectiveness of Components of Training

Bandura (1969) cited instruction, modeling, practice, and feedback as essential components of learning. With regard to counselor training, Tang et al. (2004) found that self-efficacy was moderately to strongly related to both

didactic (amount of academic coursework) and practice (number of internship hours and clinical instruction) components. Other researchers (Barbee, Scherer, & Combs, 2003; Melchert, Hays, Wiljanen, & Kolocek, 1996) have shown that trainee self-efficacy is related to the number of academic courses (didactic component) and number of clinical courses taken and hours spent with clients (practice).

In their narrative and meta-analytic review of the effects of helping skills training on the ability to learn empathy and reflections of feelings, Hill and Lent (2006) found medium to large effects for modeling (d = .90), feedback (d = .89), and instruction (d = .63). They also found that modeling was better than instruction or feedback (d = .67) and that using multiple components was better than using any single component (d = .51). Practice was not studied often enough to be included in the meta-analyses. Hill and Lent noted that researchers have not focused on training of skills other than empathic responses or reflections of feelings. Furthermore, they reported that most of the training has been brief (5- to 30-min) outside the context of training programs. This type of brief, discrete training is not appropriate for teaching insight skills such as immediacy, challenges, and interpretation, which require a foundation of exploration skills. In addition, trainees were taught to formulate simple grammatical statements (e.g., a reflection of feelings) but did not learn about the complexity of the skill, how to use the skill clinically, or how to judge the client's response. Given that skills generally are not taught in isolation, the generalization of these findings to actual training is suspect. Furthermore, the definition and the implementation of the components were vague in the previous studies. It is also difficult to determine exactly what was done in the experimental manipulations of the components, and thus it would be difficult to replicate these studies. Finally, in previous studies, the components were tested separately. It is difficult, however, to separate the components during the actual training process. How can a skill be modeled without providing some kind of definition of what is being modeled? How can realistic feedback be provided without basing it on practice? Given the nature of training, it is perhaps better to include all the components and examine their relative merits.

Predictors of Outcome of Helping Skills Training

Knowing who would be most likely to benefit from helping skills training would be of great interest to trainers. Training could be targeted specifically to meet the needs of different types of students, and students who would not benefit could be advised against going into careers involving helping skills.

Unfortunately, we have minimal empirical evidence about predictors. In their review, Hill and Lent (2006) reported no consistency of findings in five

studies examining predictors of the outcome of helping skills training because all examined different predictors (e.g., dominance, sex, conceptual level, attitudes toward the target skill, and pre-training expectations for directive vs. nondirective therapy style), making it difficult to draw any conclusions. Likewise, in the Hill et al. (2008) study, none of the predictors (initial levels of empathic concern, perspective taking, grade point average, and perfectionism) were significantly related to nine outcome variables once outliers were removed. Given the paucity of literature on predictors of success of training, further study using other predictor variables seems indicated. It also could be that predictors might be more related to outcomes in training for specific skills than for the overall effectiveness of training. Another possibility is that samples need to be larger and more diverse than those used in the past to provide enough variability in the predictors and outcomes for significant findings to emerge.

In the present series of studies, we could not rely on the previous literature for ideas of what predictors to include. Instead, we chose four variables that seemed likely to be related to an individual's ability to learn helping skills. First, given that self-efficacy for using the target skill was our primary outcome measure, we postulated that initial levels of self-efficacy would be related to both final levels and change across time. On one hand, students with higher self-efficacy might be more confident in their ability to master the skills and thus gain more from training. On the other hand, there might be a ceiling effect, such that they would not have much room to grow in terms of self-efficacy. Second, we thought that prior helping experiences might influence ability to learn the skills. On one hand, it would seem that trainees who have had more experiences would be more familiar with the concepts and thus would have an advantage in terms of growth. On the other hand, we could speculate that they would be less amenable to learning a new way of approaching helping and to receiving feedback. The third variable we thought might be important is attitude toward learning helping skills. We thought that students who are eager and open to learning the skills might learn more. Finally, we wondered if natural helping ability would be related to ability to learn helping skills. We speculated that trainees who begin training with more natural helping ability might end up with higher levels but might not change as much because they would not have as much room to grow.

Goals for the Three Studies in This Series

We had three overall goals for each of the three studies in this current series. First, we wanted to test the overall effectiveness of training for insight skills (immediacy, challenges, and interpretation). The primary outcome measure

was self-efficacy given that it has been used frequently in counseling research as an index of counselor growth (note that additional outcome measures were used within the different studies). Our second goal was to assess the effectiveness of components of training, primarily focusing on instruction, modeling, and practice given that these have been identified in past theory and research (see Hill & Lent, 2006). Note that we did not include feedback as a component because we could not determine how to standardize feedback within the context of large classes. The third goal was to examine four predictors (initial levels of self-efficacy, prior helping experiences, attitudes toward learning helping skills, and natural helping ability) of the effectiveness of training to address the question about whether some students profit more than others from training on the insight skills.

Methodological Considerations for the Three Studies in This Series

Because of the multitude of methodological problems noted earlier, we hoped to improve somewhat on the methodological rigor in these three studies. Because the present studies were conducted within the confines of regular classroom instruction, however, we were painfully aware of trade-offs that were necessary, given that instructional needs are always more important than research considerations. In addition, we are humbled by remembering Gelso's (1979) bubble hypothesis, whereby when you solve some methodological problems, others inevitably emerge. Although the current studies are not problem-free given the difficulties of conducting research in a classroom setting, we believe that they represent an improvement over the past literature. We describe here our thinking about some of the methodological issues involved in these studies.

Focus on Insight Skills

In the previous literature, as noted previously, the focus typically has been on empathy or exploration skills (e.g., reflection of feelings). Although the focus on empathy and exploration skills is certainly important as a foundation for effective listening and helping, we also need to focus on the effectiveness of training for advanced skills commonly used in psychodynamic treatments (e.g., immediacy, challenges, interpretation). We note that focusing on exploration is relatively easy because these skills are beginning or foundational skills and thus brief training programs can be employed. In contrast, insight skills typically require a foundation of ability to use the exploration skills, a theoretical introduction to the rationale for using the skills, and time for

practice. Hence, in the studies in this section (Chui et al., 2014; Jackson et al., 2014; Spangler et al., 2014), we focused on the insight skills of immediacy, challenge, and interpretation, respectively, as taught in the Hill (2004, 2009, 2014) model.

Immediacy, challenge, and interpretation are all skills in the insight stage. All three build on the foundation of exploration skills, empathy, collaboration, and self-awareness. The insight skills provide the client with a new and different perspective in a way that does not damage but rather strengthens the therapeutic relationship. There are also differences, however, among the three skills. Immediacy involves a very personal communication about the therapeutic relationship. Challenges involve addressing some kind of discrepancy in a client's viewpoints or encouraging the client to think or feel in a new way. Interpretation requires thinking in a complex way about client dynamics and underlying motivations, which requires knowledge of theories about psychopathology and psychotherapy. Thus, we expected that we would find similar results across studies but also that there might be some differences in the processes and outcomes in the three studies given that different skills were being investigated. Furthermore, we expected that there might be individual differences in how trainees perceive and respond to training for the three different skills.

Use of an Undergraduate Sample

Our decision to focus on undergraduates as our sample was based on a number of considerations. First, the initial exposure to helping skills for many students is at the undergraduate level, either through a helping skills course, training for a crisis hotline, or training to be a resident assistant in a dormitory. Although we do not have a complete estimate of how many helping skills classes are taught at the undergraduate level in the United States, a preliminary Google search (we stopped after four pages of 3,900,000 results; search terms: undergraduate program, helping skills, counseling skills) revealed 13 undergraduate programs outside of our own that advertised a helping skills course. A query to the Society of Counseling Psychology (Division 17 of the American Psychological Association) listserv revealed another 16 institutions with undergraduate helping skills courses. At our large public university, there are typically three to six undergraduate helping skills courses of at least 30 to 50 students per semester. Given that helping skills are being taught at the undergraduate level, research is needed on the effects of such training.

Second, previous studies (e.g., Hill & Kellems, 2002; Hill et al., 2008) have shown that undergraduates can learn helping skills using the Hill (2004, 2009) model. Similarly, in a meta-analysis of studies involving MC cited

earlier, Baker and Daniels (1989) found a large effect size for undergraduates (1.18) and a moderate effect size (0.66) for graduate students. In their reviews, Goodyear and Guzzardo (2000) and Hill and Knox (2013) suggested that larger changes would be expected in undergraduate than in graduate students because undergraduates have more room to grow than do graduate students who have already often had a fair amount of training.

We also argue that one way to teach a complex helping skill is to begin with basic components at an early stage of training. Just as learning basic statistics (e.g., calculating standard deviation, conducting a *t* test) at the undergraduate level provides a foundation for learning more complex statistical analyses in graduate programs, learning the rudiments of the insight skills in an undergraduate course lays the groundwork for more detailed training at the graduate level. We in no way claim that undergraduate students can master these skills during one semester, but we were interested in whether they could learn basic knowledge, awareness, and skills. We would also suggest that training in insight skills can be valuable to give undergraduate students a glimpse of what helping entails. If they have a hands-on experience of helping, they might be able to make good decisions about pursuing helping careers in the future.

A final consideration relates to sample size. Ridley et al. (2011) noted that most previous studies of helping skills training used very small sample sizes. They recommended multi-site studies as a way of increasing the sample size. We considered a multi-site study but ultimately concluded that differences in course content would make it difficult to do such a study for insight skills. Using an undergraduate sample with all instructors trained by the same person and all following the same syllabus, however, allowed us access to a larger number of potential participants and still allowed for considerable control over extraneous variables.

The use of undergraduate students, of course, raises the question about whether we can generalize the findings to graduate student training. On one hand, we would argue that students in the undergraduate courses at our university are in their final semester of undergraduate school (with a modal age of 22 years) and some go on to graduate school the following semester, which makes them similar to beginning graduate students learning helping skills. On the other hand, we would note that undergraduate students have a more diverse set of career paths than do counseling psychology graduate students; our psychology undergraduate majors go on to careers in not only psychology but also social work, law, criminal justice, and medicine. My (first author) sense from having taught both levels for many years is that undergraduates are more diverse in many ways (personality, chosen profession, world assumptions) than graduate students who tend to have a more

sophisticated understanding and acceptance of psychological constructs. If anything, training more psychologically naïve undergraduates is more challenging than training graduate students who already have often had considerable training and experience.

Design Considerations

Past studies have been criticized because they lacked random assignment, did not include control conditions, and involved only one instructor. We struggled considerably with these issues given the naturalistic nature of this research.

Given that these classes were taught within an undergraduate curriculum, students were able to choose whether or not to take the course (although they were required to take a lab course, there were several to choose from), and, within the confines of availability, they were able to choose their instructor. Because this course was very popular, enrollment often closed quickly, which meant that, because seniors had priority for registration, they were usually the only students able to take the course. Furthermore, given that this was an undergraduate course within the established curriculum, random assignment to instructor was not possible.

We thought about comparing the helping skills course with other courses. The obvious comparison would be the Introduction to Counseling course, also taught at our university by counseling psychology instructors. However, instructors of the Introduction to Counseling courses have been understandably reluctant to implement the research tasks because they make no sense for their curriculum and would not be directly relevant for their students (e.g., What rationale would instructors give for making students complete self-efficacy measures several times within the space of two classes not focused on self-efficacy for helping skills?). Furthermore, there would still also be a lack of random assignment to the two courses, and often there is only one instructor of the Introduction to Counseling course, so it would not be a good control even if we could get cooperation.

Given the constraints against having a viable control or comparison condition, we instead implemented a delay condition in the Spangler et al. (2014) and Jackson et al. (2014) studies. Half of the classes in each study received training 1 week later (delay condition) than the other half of the classes (nondelay condition), thus allowing us to assess whether changes that occurred were related to the training. We were not able to assign half of the classes to a delay condition in the Chui et al. study because of last-minute scheduling conflicts due to a snow emergency, but all participants completed the self-efficacy measure twice before training, so essentially all students served as their own control in terms of a delay.

The ideal design would be to assess the trainee's ability to use the target skill with an actual client in a naturalistic setting. This strategy is fine when assessing exploration skills (as in Hill et al., 2008) because therapists can appropriately exhibit empathy and exploration skills in a brief session. The problem with assessing the ability to use insight skills, however, is that insight skills often are not appropriate to use until later in therapy and it would be unethical to allow undergraduate students to see clients in real therapy given their lack of training. In addition, the logistics of having more than 100 trainees seeing actual clients for more than a single session within the course of a semester makes this option not feasible.

Finally, we wanted to have more than one instructor in each study, so that the results would not be confounded between the training and the instructor. Luckily, we were able to have four to five instructors in each study.

Choice of Dependent Variables (Outcomes of Training)

Given the conceptual, practical, and interpersonal challenges of teaching and learning the insight skills, it makes sense that trainees' beliefs about their ability to appropriately and effectively use these advanced skills is an important outcome of training. Counselor self-efficacy, defined as a counselor's confidence in his or her ability to use specific skills to counsel effectively in the near future (Larson & Daniels, 1998), has been recognized as essential to the development and performance of counseling skills. Bandura (1986) posited that because self-efficacy affects cognitive and emotional responses to events, as well as effort and persistence in a given domain, it is essential to determining an individual's action within that domain. In terms of how self-efficacy applies to helping skills training, trainees' beliefs about being able to use helping skills influence their choice of skill at a given moment, how much effort they choose to expend in using skills correctly, how persistent they are when faced with difficulty, and how well they are able to manage their cognitive and emotional state. Relatedly, Lent, Hill, and Hoffman (2003) and Lent et al. (2006) indicated that graduate trainees' counseling self-efficacy increased over the course of their first practicum experience.

Bandura (1991) also posited that self-efficacy is reinforced by successful performance of tasks, thus further building personal agency and enabling the individual to respond to and shape events in his or her environment. We note that although theoretically the relationship between self-efficacy and performance should be high, in fact the results are mixed. For example, Larson and colleagues (1992) found that trainee counseling self-efficacy was significantly correlated with performance on specific counseling microskills. Similarly, Lent et al. (2009) indicated that trainees' beliefs about their own

counseling ability were related to how they responded to clients. Heppner, Multon, Gysbers, Ellis, and Zook (1998) found, however, that although both trainee ratings of self-efficacy for career counseling increased from pre-practicum to post-practicum and client ratings across several career counseling outcomes increased from pre- to post-test, the relationship between self-efficacy and career counseling process and outcome was mixed at best. In addition, when Lent et al. (2006) used a client-specific measure of self-efficacy, they found associations between self-efficacy and trainee ratings of session outcome but not with client ratings of session outcome. Hence, a limitation of the use of self-efficacy as a measure of training outcome is that it may not be related to performance of skills or session outcome.

For this series of studies then, our primary assessment of outcome involved self-report measures of self-efficacy. Previous research on the Hill training model (Hill et al., 2008) included an assessment of self-efficacy, but this assessment involved only a single item. Given the psychometric problems with using single-item measures and the notion that self-efficacy is domain-specific (i.e., related to the specific topic), the development of reliable multi-item measures of self-efficacy specific to each target skill was needed. Hence, we created separate self-efficacy measures for immediacy, challenge, and interpretation. Furthermore, because students were completing the measure after being exposed to each component of training and thus would complete the measure up to 10 times, we had to create measures that were not too cumbersome and could be completed quickly during class time. Thus, each of the measures had only four items. We recognize that reactivity could be a problem given the number of times completed.

In addition to self-efficacy, we used other measures of training outcome. In the study on training for immediacy (Spangler et al., 2014), we included the Catharsis and the Cohesion subscales from the Curative Climate Instrument (Fuhriman, Drescher, Hanson, Henrie, & Rybicki, 1986). Although the Catharsis subscale is intended to be a measure of the helpfulness and frequency of expressing emotionally charged material, it includes items that assess how helpful it is to be able to speak forthrightly about feelings about oneself in relation to others in the group in the moment, that is, immediacy. Thus, if group members perceive this factor as helpful, we would have some evidence of how well immediacy is being done. We also speculated that lab group members' perceived cohesion of their group would increase as a result of their learning to be immediate with each other. In the study on training for challenges (Chui et al., 2014), we included an assessment of quality of written challenges provided in response to written client stimuli, even though we recognized the limitations of such measurement (as described earlier in this article). Finally, in the Jackson et al. (2014) study on

training for interpretation, we used a slightly different procedure, such that the students who were role-playing clients were given a script to act out that set them up for a possible interpretation, and students who were the helpers were required to give at least one interpretation. After each 20-min session, helpers and clients completed two items from the Helping Skills Measure (Hill & Kellems, 2002) assessing the perceived use of interpretations.

Components of Training

As noted previously, we designed our training around Bandura's (1969, 1986, 1989) components of instruction, modeling, practice, and feedback. We divided instruction into reading and lecture to determine if there were differences in these two methods. Modeling typically involved some video demonstrations by expert therapists. Practice included written responses to written client stimuli, large group practice, and practice in dyads. Because of the large class sizes, it was not possible to specifically manipulate feedback, but we allowed instructors and lab leaders to provide feedback as they normally would (training as usual), suggesting that feedback was intertwined with practice. We used only one sequence for the order of components: reading about the skill, listening to a lecture, watching a model, and then practicing with feedback using a variety of practice exercises. From a research design perspective, testing for order effects would be optimal, but from a training perspective, it makes sense to have reading followed by lecture, and then modeling, and finally practice for all trainees. Each of these components builds on the other, such that it would not be pedagogically sound to practice the skill before learning about it through instruction and modeling.

The effects of each of the components were tested through changes in self-efficacy for using that particular skill. Hence, after the presentation of each component, students completed the self-efficacy measure. We also included a qualitative measurement strategy that allowed us to hear from students about other possible aspects of the training that might have been helpful or hindering. Specifically, in the Spangler et al. (2014) study, we asked the students to write a paper immediately after the training, reflecting on difficulties in training and the most and least helpful aspects of training. The feedback from the qualitative analyses was then used in designing the Chui et al. (2014) study. For example, based on the feedback, we included more and better opportunities for practice and better videos for modeling in the Chui et al. study. In addition, based on the results, we developed rating scales for students to use to evaluate the various components of training, although we still asked for a narrative description of their ratings. Because the findings from the qualitative analyses of the first two studies were relatively consistent, we

did not include a qualitative component in the Jackson et al. (2014) study, but we did include similar rating scales of the components as we used in the Chui et al. study. The self-efficacy measures reflected change immediately after the presentation of each component, whereas the narratives and ratings were completed at the end of training when students could reflect on the relative helpfulness and effects of all of the components.

Methods for the Three Studies

The three studies were conducted sequentially. We collected data for the Spangler et al. (2014) study on training immediacy in Spring semester 2009. Based on the results, we made some modifications to the design and then collected data for the Chui et al. (2014) study on training challenges in Spring semester 2010. Finally, we again made modifications to the design and collected data for the Jackson et al. (2014) study on training interpretation in Spring semester 2011. In this section, we describe methods used in all three studies. Methods specific to each of the studies are described in the individual papers.

The Helping Skills Course

The 15-week, four-credit psychology laboratory course focused on teaching skills based on the Hill (2004, 2009) three-stage model of helping. This undergraduate course was one of five that fulfilled the departmental requirement of two laboratory classes. The course objectives were for students to gain knowledge of and practice using helping skills, examine the major theoretical approaches to helping skills, learn to evaluate research on helping skills, and gain knowledge about special topics in counseling (e.g., intake, termination). The course involved a 2-hr weekly lecture/discussion with about 33 students, as well as a 2-hr weekly laboratory section in which students practiced skills in small groups of 8 to 12 students. Students engaged in one 20-min helping session in the first or second week of the semester, and then another 20-min session two thirds of the way through the semester. For these sessions, students were paired with classmates (a different person in each session). One student served as the helper, and the other as a client; after the first session, they switched roles and conducted another session. For course requirements, students transcribed the sessions in which they were helpers and then coded the skills used as well as indicated what skills they wish they had used and coded these. These two transcripts served as the basis for a lab report in which they evaluated their skills and their progress in learning the skills. Students also wrote a self-examination paper at the beginning of the semester, and took midterm and final examinations.

Students were told that, in accordance with ethical guidelines, they were not required to disclose personal information and could role-play problems during practice if they preferred. Although we do not have empirical data to support this claim, most students indicated (and our observations confirmed) that they discussed real but not deeply disturbing or inappropriate problems (e.g., roommate concerns, relationship difficulties, academic stresses, and career plans) during practice sessions.

Measures Used in All Three Studies

Demographic measure. A demographic questionnaire asked students' age, gender, race/ethnicity, major, and previous coursework in counseling/clinical psychology.

Self-efficacy measures. Because self-efficacy is domain-specific, separate but similar four-item measures were created for each study to assess the helper's self-efficacy for using the target skill. In all three studies, trainees rated their confidence in their ability to use immediacy on a 10-point scale, ranging from 0 (*no confidence*) to 9 (*complete confidence*). Scores for all measures were an average of the four items.

The *Self-Efficacy for Immediacy* (SEIm) items are as follows: "I can use immediacy in a session with a client," "I can use immediacy to talk in the here-and-now to a client about our relationship," "I can talk in the here-and-now about positive aspects of my relationship with my client," and "I use immediacy to address problems or misunderstandings between us as they arise." In principal-axis factor analyses of the seven administrations of the SEIm, the KMO index was satisfactory (.82–.85), and the Bartlett chi-squares ($df = 6$, range = 356.38–533.30) were significant at the $p < .001$ level. There was always only one eigenvalue >1 (range = 3.24–3.50) and the scree plots suggested a single factor accounting for 81% to 86% of the variance; all items loaded greater than .78, and internal consistency alphas ranged from .92 to .95.

The *Self-Efficacy for Challenge* (SEC) items are as follows: "I can use challenges in a session with a client," "I can point out discrepancies in what a client says," "I can point out contradictions between a client's words and behaviors," and "I use challenges to point out a client's maladaptive thoughts." In principal-axis factor analyses on the 10 administrations of the SEC, the KMO index was satisfactory (.66–.86), and the Bartlett chi-squares ($df = 6$, range = 213.85–477.51) were significant at the $p < .001$ level. There was always only one eigenvalue >1 (range = 2.79–3.48), the scree plots always suggested a single factor accounting for 70% to 87% of the total variance,

and all items loaded greater than .73 on the single factor. Internal consistency alphas ranged from .84 to .95.

The *Self-Efficacy for Interpretation* (SEIn) items are as follows: "I can use interpretations in a session with a client," "I can point out patterns or themes in what my client is saying," "I can help a client gain a deeper understanding of what he or she discusses in a helping session," and "I can provide the client with new meaning or a new explanation regarding his or her behaviors, thoughts, or feelings." In principal-axis factor analyses on the 11 administrations of the SEIn, the KMO index was satisfactory (.80–.86), and the Bartlett chi-squares ($df = 6$, range = 171.44–576.79) were significant at the $p < .001$ level. There was always only one eigenvalue >1 (range = 2.68–3.54), the scree plots always suggested a single factor accounting for 67% to 88% of the total variance, and all items loaded greater than .78 on the single factor. Internal consistency alphas ranged from .83 to .96.

Prior Helping Experiences (PHE). This self-report measure was created by the authors to assess the amount of experience the students had helping others prior to the course. Two items asked about the level of experience (e.g., "Completed coursework in peer counseling, peer mediation, or helping skills") and used a 5-point scale (0 = *none at all*, 4 = *extensive*). Two items asked about practical experience with five options measuring increasing levels of experience (e.g., "I have had some experience prior to college helping clients directly, such as peer mediation in high school"). For the latter two items, a check on the first option was rated 1, whereas checks on subsequent options were rated 2, 3, 4, and 5, respectively. A principal-axis factor analysis in Spangler et al. (2014) revealed that the KMO index was satisfactory (.68), and the Bartlett chi-square ($df = 6$, 251.73) was significant at the $p < .001$ level; there was one eigenvalue >1 (2.21) and the scree plot suggested a single factor accounting for 66% of the variance; all items loaded greater than .59. Internal consistency (alpha) ranged from .82 to .86 in the three studies. The score was an average of the four items, with higher scores indicating more prior experiences.

Attitudes toward Learning Helping Skills (ALHS). This self-report measure was created by the authors to assess motivation for learning the skills. Four items (e.g., "I am eager to learn helping skills") are rated on a 9-point scale (1 = *completely disagree*, 5 = *neutral*, 9 = *completely agree*). A principal-axis factor analysis of the data from the Spangler et al. (2014) study revealed that the KMO index was satisfactory (.72) and the Bartlett chi-square ($df = 6$, 317.40) was significant at the $p < .001$ level. There was one eigenvalue >1 (2.78), and the scree plot suggested a single factor accounting for 69% of the variance;

all items loaded greater than .58. Internal consistency (alpha) ranged from .84 to .86 in the three studies. The score was an average of the four items, with higher scores indicating more positive attitudes.

Natural Helping Measure (NHM). This self-report measure has five items (e.g., "I often find myself helping others with their problems") using a 7-point Likert-type scale from 1 (*never*) to 7 (*always*) and assesses inherent helping ability. In a factor analysis, Stahl and Hill (2008) found a one-factor structure accounting for 51% of the variance; all items loaded >.50. Test–retest reliability over 2 to 4 weeks was .67; internal consistency was .81. Internal consistency (alpha) ranged from .84 to .87 in the three studies. The score was an average of the five items, with higher scores indicating more natural helping ability.

Procedures

To minimize variation among courses, the instructors in each study met before the start of the semester to discuss their syllabi. They planned to cover topics in a similar sequence and agreed on policies regarding assignments. Although we have no assessment of adherence to a specific training protocol, all (other than the first author) were advanced students in the same counseling psychology doctoral program, had been trained in helping skills by the first author, and valued helping skills and helping skills training. Prior to the semester, all instructors, teaching assistants, and lab leaders signed a consent form and completed a demographic form. Instructors and lab leaders were trained in the appropriate procedures before the section of the course on the skill.

Recruiting participants and pre-tests. During the first or second class of the semester, a researcher not associated with the course explained that completion of all the measures and helping sessions was required for the course but that students had the option as to whether to contribute their data to the study and that they would receive extra credit for participation (other options for gaining extra credit were also provided). During this same class, students signed consent forms and then completed pre-semester measures (demographics, ALHS, NHM, and PHE). To protect confidentiality, students used their university ID numbers on all measures, and instructors were informed about who participated so that they could assign extra credit only after all grading was done.

Data collection. Students completed a self-efficacy measure for the target skill at the end of the class prior to the training on the target skill. They were then told to read the relevant chapter in the textbook (Hill, 2004, 2009) and

reminded that they would be quizzed on the reading. At the beginning of the next lecture class, students took the self-efficacy measure to test the effects of reading. They then took a quiz, which all of the students passed (indicating to us that they had done the reading). After instructors gave an approximately 30-min lecture about the skill, students took the next self-efficacy measure. Instructors then presented examples (models) of therapists using the target skill (to be described in each study), after which the students completed the self-efficacy measure. Next, different types of practice were implemented in the lecture class, followed by another completion of the self-efficacy measure. At the beginning of the lab class, students again first completed the self-efficacy measure. They then participated in various practice exercises, again followed by the administration of the self-efficacy measure.

Quantitative Analyses Used Across the Three Studies

Because we were interested in change in self-efficacy over time, we conducted growth curve analyses using Hierarchical Linear and Nonlinear Modeling (HLM, Version 7.0 Student; Raudenbush, Bryk, Cheong, Congdon, & du Tolt, 2011). Students were nested within courses (instructors). Unfortunately, although we had more than 100 students in each of the studies, we had only four or five instructors in each study and thus did not have sufficient power to detect instructor effects.

The first model in the Spangler et al. (2014) and Jackson et al. (2014) studies tested for the overall effectiveness of training by analyzing differences in the delay and nondelay conditions. To test these differences, we constructed a linear (constant rate of change), a quadratic (with one peak or trough), and a cubic (with one peak and one trough) unconditional growth model because results of the graphed means of self-efficacy for each target skill indicated the possibility of a nonlinear, cubic pattern of change during training. For the unconditional models, the data were centered on Time 1. For brevity, we include here only the unconditional cubic growth the Level 1 (within-person change over time) model:

$$Y_{ti} = \pi_{00} + \pi_{10}\left(\text{Time}\right)_{ti} + \pi_{20}\left(\text{Time}\right)_{ti}^{2} + \pi_{30}\left(\text{Time}\right)_{ti}^{3} + r_{0i}, \qquad (1)$$

where Y_{ti} is individual i's self-efficacy score at time t; π_{00} is the intercept, which is the individual's self-efficacy score at the start of training; $\pi_{10}(\text{Time})_{ti}$ represents the linear rate of change in self-efficacy for the individual;

$\pi_{20}(\text{Time})_{it}^2$ represents the quadratic rate of change; $\pi_{30}(\text{Time})_{it}^3$ represents the cubic rate of change; and r_{0i} represents error. This level examined within-person difference in self-efficacy. The Level 2 (between-person differences in intercept and slope) unconditional cubic growth model was as follows:

$$\begin{aligned}\pi_{0i} &= \beta_{00} + r_0\left(\text{error}\right),\\ \pi_{1i} &= \beta_{10} + r_1\left(\text{error}\right),\\ \pi_{2i} &= \beta_{20} + r_2\left(\text{error}\right),\\ \pi_{3i} &= \beta_{30} + r_3\left(\text{error}\right),\end{aligned} \tag{2}$$

where β_{00} represents the overall mean initial self-efficacy for all participants, β_{10} is the overall mean linear rate of change in self-efficacy for all participants, β_{20} is the overall mean quadratic rate of change, β_{30} is the overall cubic rate of change, and r_0, r_1, r_2, and r_3 are error. This level examined the between-person difference in self-efficacy for the target skill.

To determine which of the significant unconditional models explained the most within-person variance, pseudo-R^2 (Singer & Willett, 2003) was calculated for each model using the sigma-squared statistics for the unconditional means model and for the significant linear and cubic unconditional growth models as follows:

$$\frac{\sigma^2\left(\text{unconditional means}\right) - \sigma^2\left(\text{unconditional growth}\right)}{\sigma^2\left(\text{unconditional means}\right)}. \tag{3}$$

For the Spangler et al. (2014) and Jackson et al. (2014) studies, the unconditional cubic growth model explained a greater proportion of within-person change, and there was a significant variance in the unconditional cubic model (results for self-efficacy for the target skills are presented in the individual articles). Thus, we proceeded to test a conditional cubic model. For this model, a parameter for the nondelay/delay condition was added at Level 2. The Level 1 data were re-centered on the time point at which participants in the nondelay condition had completed the training on the target skill and those in the delay condition had not yet begun the training. The Level 1 model was the same as Equation 1. At Level 2, we added the following nondelay/delay condition:

$$\begin{aligned}\pi_{0i} &= \beta_{00} + \beta_{01} + r_0,\\ \pi_{1i} &= \beta_{10} + \beta_{11} + r_1,\\ \pi_{2i} &= \beta_{20} + \beta_{21} + r_2,\\ \pi_{3i} &= \beta_{30} + \beta_{31} + r_3,\end{aligned} \tag{4}$$

where β_{00} represents the overall mean initial self-efficacy for all participants; β_{10}, β_{20}, and β_{30} are the overall mean linear, quadratic, and cubic rates of change in self-efficacy for all participants; β_{01}, β_{11}, β_{21}, and β_{31} are the delay versus nondelay parameter for the mean and linear, quadratic, and cubic slopes, respectively; and r_0, r_1, r_2, and r_3 are error. This level examined the between-person variance in self-efficacy and differences between the delay and nondelay groups. Coefficients, standard errors, and t ratios for fixed effects and variance and chi-square values for random effects are presented in the individual articles. For the Chui et al. (2014) study, only the linear growth model was significant. Because there were no delay/nondelay conditions in the Chui et al. study, the Level 1 data remained centered on Time 1.

Predictors of Final Levels of, and Change in, Self-Efficacy for Target Insight Skill

We first examined the bivariate correlations between the predictor variables and the outcome variables (see individual articles). Based on the correlation results, we created growth curve models with the significantly correlated variables added at Level 2. For this model, because we were not testing the effects of nondelay–delay, the data for delay and nondelay conditions were combined along the same timeline, so that for all participants, Self-Efficacy-1 was pre-training, Self-Efficacy-2 was post-reading, and so forth. For the unconditional models, the data were centered on Self-Efficacy-1.

Fixed effects and variance components for the unconditional means and growth models for all three studies were significant. Given the unexplained variance in the unconditional models, we proceeded with conditional models that included initial self-efficacy and PHE as predictors at Level 2 for the immediacy and interpretation studies (none of the predictors were significantly correlated with the change in the challenge study and so no conditional models were generated). Once we determined the effects of the delay versus nondelay groups, we combined the groups along the same timeline to test the predictors. We also re-centered the data on the final self-efficacy time point for each study. We then tested the combined, re-centered data and found no difference between nondelay and delay on final self-efficacy or on the linear, quadratic, or cubic components. We then constructed the predictor models. For the immediacy study, the Level 1 model was the same as Equation 1, and the Level 2 conditional growth model was as follows:

$$\begin{aligned}\pi_{0i} &= \beta_{00} + \beta_{01} + \beta_{02} + r_0\,(\text{error}),\\ \pi_{1i} &= \beta_{10} + \beta_{11} + \beta_{12} + r_1\,(\text{error}),\\ \pi_{2i} &= \beta_{20} + \beta_{21} + \beta_{22} + r_2\,(\text{error}),\\ \pi_{3i} &= \beta_{30} + \beta_{31} + \beta_{32} + r_3\,(\text{error}),\end{aligned} \tag{5}$$

where β_{00} represents the overall mean initial self-efficacy for all participants; β_{10}, β_{20}, and β_{30} are the overall mean linear, quadratic, and cubic rates of change in self-efficacy for all participants; β_{01}, β_{11}, β_{21}, and β_{31} are initial self-efficacy for the mean and linear, quadratic, and cubic slopes, respectively; β_{02}, β_{12}, β_{22}, and β_{32} are PHE for the mean and linear, quadratic, and cubic slopes, respectively; and r_0, r_1, r_2, and r_3 are error. This level examined whether initial self-efficacy and PHE predicted between-person variance in self-efficacy.

The cubic model was significant for the Spangler et al. (2014) immediacy study; however, a linear model was a better fit for the Jackson et al. (2014) interpretation study. Thus, the Level 1 growth model for those studies was the same as Equation 1. For the interpretation training study, the Level 2 growth model was as follows:

$$\begin{aligned}\pi_{0i} &= \beta_{00} + \beta_{01} + \beta_{02} + r_0,\\ \pi_{1i} &= \beta_{10} + \beta_{11} + \beta_{12} + r_1,\end{aligned} \tag{6}$$

where β_{00} represents the overall mean initial self-efficacy for all participants; β_{10} is the overall mean linear rate of change in self-efficacy for all participants; β_{01} and β_{11} are initial self-efficacy for the mean and linear slopes, respectively; β_{02} and β_{12} are PHEs for the mean and linear slopes, respectively; and r_0 and r_1 are error. This level examined whether initial self-efficacy and PHE predicted between-person variance in self-efficacy.

The coefficients, standard errors, and t ratios for fixed effects, as well as variance and chi-square values for random effects for models appear in the Results sections of the individual articles.

Effects of Training Components Assessed Quantitatively

For these analyses, the data for participants in the delay and nondelay conditions were combined along the same timeline. Repeated-measures ANOVAs followed by pairwise comparisons were used to investigate the unique contribution of each component of training to the change in self-efficacy for each skill.

Qualitative Analyses of the Narratives in Spangler et al. (2014) and Chui et al. (2014)

To analyze the narratives about their experiences of the training, we used modified consensual qualitative research (CQR-M; Spangler, Liu, & Hill, 2012), which allowed us to analyze relatively limited data from the large samples (students did not write much in response to each question, so it was not necessary to do the more extensive CQR analyses typically required for interview data). The first two authors for each of the studies read about 30 reflection papers and created categories for domains (e.g., most helpful components of training, least helpful components of training, difficulties involved in using the skill, and cultural influences on the ability to learn the skill) that reflected what the students wrote. The first two authors then met with other authors and modified the category list by reviewing another 30 papers. Then, at least two authors consensually coded each thought unit in the remaining reflection papers into one or more categories. Note that narratives were not collected in the Jackson et al. (2014) study.

Conclusion

In this article, we have presented an overview of the rationale, methods, and analyses for the three studies that follow in this series of articles. Following the three studies, we present a final article in which we summarize and discuss the findings across studies.

Authors' Note

Patricia Spangler is now at the Center for the Study of Traumatic Stress, Department of Psychiatry, Uniformed Services University of the Health Sciences, Bethesda, MD. Harold Chui is now on a postdoctoral fellowship at Adelphi University, Garden City, NY. John Jackson is now at the Wellness Center at Sewanee: The University of the South, Sewanee, TN.

Declaration of Conflicting Interests

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

Funding

The author(s) received no financial support for the research, authorship, and/or publication of this article.

References

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

Bandura, A. (1969). *Principles of behavior modification*. New York, NY: Holt, Rinehart and Winston.

Bandura, A. (1986). Self-efficacy. In A. Bandura (Ed.), *Social foundations of thought and action: A social cognitive theory* (pp. 390-453). Englewood Cliffs, NJ: Prentice Hall.

Bandura, A. (1989). Human agency in social cognitive theory. *American Psychologist*, *44*, 1175-1184. doi:10.1037/0003-066X.44.9.1175

Bandura, A. (1991). Human agency: The rhetoric and the reality. *American Psychologist*, *46*, 157-162. doi:10.1037/0003-066X.46.2.155

Barbee, P. W., Scherer, D., & Combs, D. C. (2003). Prepracticum service learning: Examining the relationship with counselor self-efficacy and anxiety. *Counselor Education and Supervision*, *43*, 108-119.

Carkhuff, R. R. (1969). *Human and helping relations* (Vols. 1 & 2). New York, NY: Holt, Rinehart and Winston.

Chui, H., Hill, C. E., Ain, S., Ericson, S. K., Ganginis Del Pino, H. V., Hummel, A. M., . . .Spangler, P. T. (2014). Training undergraduate students to use challenges. *The Counseling Psychologist, 42*, 758-777.

Ford, J. D. (1979). Research on training therapists and clinicians. *Review of Educational Research*, *49*, 87-130.

Fuhriman, A., Drescher, S., Hanson, E., Henrie, R., & Rybicki, W. (1986). Refining the measurement of curativeness: An empirical approach. *Small Group Behavior*, *17*, 186-201. doi:10.1177/104649648601700204

Gelso, C. J. (1979). Research in counseling psychology: Methodological and professional issues. *The Counseling Psychologist*, *8*(3), 7-35. doi:10.1177/001100007900800303

Goodyear, R. K., & Guzzardo, C. R. (2000). Psychotherapy supervision and training. In S. Brown & R. W. Lent (Eds.), *Handbook of counseling psychology* (3rd ed., pp. 83-108). New York, NY: John Wiley.

Heppner, M. J., Multon, K. D., Gysbers, N. C., Ellis, C. A., & Zook, C. E. (1998). The relationship of trainee self-efficacy to the process and outcome of career counseling. *Journal of Counseling Psychology*, *45*, 393-402.

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: John 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

Ivey, A. E. (1971). *Microcounseling: Innovations in interviewing training*. Springfield, IL: Charles C. Thomas.

Jackson, J. L., Hill, C. E., Spangler, P. T., Ericson, S. K., Merson, E. S., Liu, J., . . .Reen, G. (2014). Training undergraduate students to use interpretation. *The Counseling Psychologist*, *42*, 778-799.

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: Charles C. 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. doi:10.1037/0033-2909.85.3.467

Larson, L. M., & Daniels, J. A. (1998). Review of the counseling self-efficacy literature. *The Counseling Psychologist*, *26*, 179-218. doi:10.1177/0011000098262001

Larson, L. M., Suzuki, L. A., Gillespie, N., Potenza, M. T., Bechtel, M. A., & Toulouse, A. L. (1992). Development and validation of the Counseling Self-Estimate Inventory. *Journal of Counseling Psychology*, *39*, 105-120. doi:10.1037/0022-0167.39.1.105

Lent, R. W., Cinamon, R. G., Bryan, N. A., Jezzi, M. M., Martin, H. M., & Lim, R. (2009). Perceived sources of change in trainees' self-efficacy beliefs. *Psychotherapy: Theory, Research, Practice, Training*, *46*, 317-327. doi:10.1037/a0017029

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 counseling: Relation to perceptions of session strength. *Journal of Counseling Psychology*, *53*, 453-463.

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, NY: John 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, NY: John 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, NY: John Wiley.

Melchert, T. P., Hays, V. L., Wiljanen, L. M., & Kolocek, A. K. (1996). Testing models of counselor development with a measure of counseling self-efficacy. *Journal of Counseling & Development*, *74*, 640-644.

Raudenbush, S. W., Bryk, A. S., Cheong, Y., Congdon, R., & du Toit, M. (2011). *HLM 7.0: Hierarchical Linear Modeling (student)*. Lincolnwood, IL: Scientific Software International.

Ridley, C. R., Mollen, D., & Kelly, S. M. (2011). Microskills training: Evolution, reexamination, and call for reform. *The Counseling Psychologist*, *39*, 800-824.

Rogers, C. R. (1942). *Counseling and psychotherapy*. Boston, MA: Houghton Mifflin.

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, NY: John Wiley.

Singer, J. D., & Willett, J. B. (2003). *Applied longitudinal data analysis: Modeling change and event occurrence*. New York, NY: Oxford University Press.

Spangler, P. T., Hill, C. E., Dunn, M. G., Hummel, A. M., Walden, T. T., Liu, J., . . .Salahuddin, N. M. (2014). Training undergraduate students to use immediacy. *The Counseling Psychologist, 42*, 729-757.

Spangler, P. T., Liu, J., & Hill, C. E. (2012). CQR for simple qualitative data: An introduction to CQR-M. In C. E. Hill (Ed.), *Consensual qualitative research: A practical resource for investigating social science phenomena* (pp. 269-284). Washington, DC: American Psychological Association.

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

Tang, M., Addison, K. D., LaSure-Bryant, D., Norman, R., O'Connell, W., & Stewart-Sicking, J. A. (2004). Factors that influence self-efficacy of counseling students: An exploratory study. *Counselor Education and Supervision*, *44*, 70-80.

Author Biographies

Clara E. Hill, PhD, is a Professor of Psychology in the Counseling Psychology program at the University of Maryland. She received her doctoral degree in counseling psychology from Southern Illinois University in 1974 and has been at the Department

of Psychology at the University of Maryland since then. Her current areas of interest are the identification and training of counseling skills, process and outcome studies of psychotherapy, working with dreams, and qualitative research.

Patricia T. Spangler, PhD, is a Clinical Research Psychologist at the Center for the Study of Traumatic Stress, Department of Psychiatry, Uniformed Services University of the Health Sciences. She completed her doctoral degree in psychology at the University of Maryland in 2010. Her research interests include counseling and helping skills training, psychotherapy process and outcome, and treatment development for posttraumatic stress disorder with a focus on nightmares and sleep disturbance.

Harold Chui, PhD, is a Post-Doctoral Research Fellow at The Derner Institute of Advanced Psychological Studies at Adelphi University in Long Island, NY. He completed his doctoral degree in counseling psychology at the University of Maryland in 2014. His research interests include psychotherapist training and therapist factors in psychotherapy process and outcome.

John L. Jackson, PhD, is a Staff Psychologist at Sewanee: The University of the South's Wellness Center. He completed his doctoral degree in psychology at the University of Maryland in 2013. His research interests include counseling and helping skills training, psychotherapy process and outcome, and the application of attachment theory to the development and maintenance of the therapeutic relationship.