The Counseling Psychologist

Training Undergraduate Students to Use Insight Skills: Integrating the Results of Three Studies

Clara E. Hill, Patricia T. Spangler, John Jackson and Harold Chui

Additional services and information for *The Counseling Psychologist* can be found at:

Email Alerts: http://tcp.sagepub.com/cgi/alerts

Subscriptions: http://tcp.sagepub.com/subscriptions

Reprints: http://www.sagepub.com/journalsReprints.nav

Permissions: http://www.sagepub.com/journalsPermissions.nav

>> OnlineFirst Version of Record - Jul 18, 2014

What is This?

Training Undergraduate Students to Use Insight Skills: Integrating the Results of Three Studies

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

Abstract

We compare the results of a series of studies (Chui et al., Jackson et al., and Spangler et al.) investigating the effects of training undergraduate students in helping skills courses to use insight skills (immediacy, challenges, interpretation) after they had learned exploration skills. A comparison of students and instructors indicated similarity across the samples. Increases in self-efficacy for the target skill were found across all studies. In addition, all components (reading, lecture, video modeling, practice, and feedback) were found to be effective, but students found practice and lecture to be particularly effective in all three studies. Higher initial self-efficacy for the target skill and higher prior helping experiences were associated with higher final levels of self-efficacy; lower self-efficacy for the target skill and higher prior helping experiences were associated with greater gains in self-efficacy over the course of training in two studies. Limitations and implications across all three studies are discussed.

Keywords

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

[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

In this series of three studies, we investigated the effectiveness of training undergraduate students in helping skills courses to use insight skills. The overall rationale and methods for the studies were outlined in an overview article (Hill, Spangler, Chui, & Jackson, 2014). The findings were then reported separately for each study (Chui et al., 2014; Jackson et al., 2014; Spangler et al., 2014). In this final article, we discuss the findings across the three studies. The discussion of the findings is divided into four sections: comparison of samples, the overall effectiveness of training, the effectiveness of the components of training, and the prediction of the effects of training.

Comparison of Samples

All three studies were conducted at the same university during different semesters, so we expected the student samples to be relatively similar. Indeed, that seemed to be true. Table 1 shows that student samples were very similar in terms of age, percentage of female students, percentage of European American students, scores on the Prior Helping Skills measure, scores on the Attitudes toward Learning Helping Skills measure, scores on the Natural Helper Measure, and scores on the initial measures of self-efficacy for the target skill.

Likewise, the instructor samples were very similar across the three studies. Table 1 shows that instructors all believed strongly in the Hill helping skills model and were similar in terms of theoretical orientation. In addition, we tested for instructor effects in each of the three studies and found no such effects, suggesting that results were similar across instructors. Thus, we feel relatively confident that differences in findings among the three studies are not due to differences in the student and instructor characteristics noted previously. Of course, as noted below, these and other characteristics may well influence who benefits from training.

The Overall Effectiveness of Training

Across the three studies, the major outcome variable was self-efficacy for implementing the target skill (different, but structurally similar, four-item measures were created for each study). In addition, we used other outcome measures for each of the studies. Table 2 summarizes the effect sizes for all the outcome measures across training.

Self-Efficacy as the Outcome Measure

First, we compare self-efficacy ratings from the week before the training on the insight skill to the ratings after the final practice exercise. Self-efficacy

Table 1. Comparison of Samples Across Studies.

	Spangler et al. (2014)		Chui et al. (2014)		Jackson et al. (2014)	
	M	*SD*	*M*	*SD*	*M*	*SD*
Students						
Age	21.51	1.20	21.61	2.74	21.48	2.06
% female	73		81		74	
% European American	70		71		66	
PHE scores	2.09	0.73	1.55	0.95	1.49	0.96
ALHS scores	7.74	1.00	7.70	1.34	7.89	1.13
NHM scores	5.49	0.86	5.56	0.86	5.54	0.89
Initial SE	6.41	1.82	5.50	1.34	5.67	1.65
Instructors						
Belief in Hill model	8.60	0.55	8.50	0.58	8.40	0.89
Feminist/multicultural	3.42	0.86	3.75	0.96	3.40	0.89
Psychodynamic	3.31	1.01	4.50	0.58	4.20	0.45
Humanistic	4.35	0.63	3.50	0.58	3.60	0.89
Cognitive-behavioral	3.73	0.87	2.00	1.41	2.20	1.09

Note. PHE = Prior Helping Experiences, scores range from 0 to 4; ALHS = Attitudes toward Learning Helping Skills, scores range from 1 to 9; NHM = Natural Helper Measure, scores range from 1 to 7; SE = Self-efficacy, scores range from 0 to 9. Belief in Hill model = Belief in the Hill helping skills model, scores range from 1 to 9. Feminist/multicultural, Psychodynamic, Humanistic, and Cognitive-Behavioral refer to theoretical orientations, scores range from 1 to 5. High scores on all measures reflect higher levels of the construct.

increased from 4.61 (*SD* = 1.82) to 6.48 (*SD* = 1.27) in Spangler et al. (2014) for immediacy, 5.50 (*SD* = 1.34) to 7.24 (*SD* = 1.20) in Chui et al. (2014) for challenges, and 5.67 (*SD* = 1.65) to 6.67 (*SD* = 1.34) in Jackson et al. (2014) for interpretation. Because means are correlated, effect sizes were determined by calculating a difference score for each participant and then dividing the average difference score by the pooled average standard deviation for the sample. The effect sizes were large across all three studies (see line in Table 2 for overall effects). These data provide compelling evidence that undergraduate students increased in self-efficacy for using the insight skills after reading the text and participating in 4 hr of intensive training. It is important to emphasize that this training took place following half a semester of training in exploration skills, which set the foundation for learning the insight skills.

We also have evidence that the changes in self-efficacy were not due to time alone. In two of the studies, half of the classes were assigned to a delay condition so that we could determine whether there were changes before

Table 2. Effect Sizes for Effects of Training Across Studies.

	Spangler et al. (2014)	Chui et al. (2014)	Jackson et al. (2014)
Self-efficacy			
Baseline	—	0.10	—
Delay	0.06	—	0.21
Reading	0.01	0.19	0.00
Lecture	0.83	0.66	0.52
Video modeling	0.42	0.52	0.08
Large group practice	0.34	0.51	0.44
Lab group practice	1.05	0.36	0.50
Dyad practice	—	0.24	0.37
Overall effects	1.05	1.11	0.83
Maintenance	—	0.11	—
Catharsis	0.27	—	—
Cohesion	1.00	—	—
Quality of written challenges	—	0.71	—
Interpretation use—Client-rated	—	—	0.25
Interpretation use—Helper-rated	—	—	0.40

Note. Within-sample effect sizes were calculated by computing a difference score between variables for each person and then dividing the average difference score by the pooled average standard deviation for the sample. Reported effect sizes represent changes from before the component was presented to after the component was presented. For effect sizes, > 0.20 = small, > 0.50 = medium, > 0.80 = large.

students received training in the skill. The effect sizes for the changes across the delay condition were 0.06 for Spangler et al. (2014) and 0.21 for Jackson et al. (2014). In the Chui et al. (2014) study, no changes were found for students from the beginning of the semester to right before training in challenges ($d = 0.10$). These data provide compelling evidence that students did not gain self-efficacy during the period of time prior to training in the target skill. Furthermore, students in the Chui et al. study maintained their gains in self-efficacy for 5 weeks after training, indicating that changes persisted over a short period of time while they were still in training.

We also compared across samples in terms of self-efficacy ratings using regular effect size calculations (difference between means divided by the pooled standard deviation). Students in the Spangler et al. (2014) study started with lower self-efficacy than did students in the Chui et al. (2014) and Jackson et al. (2014) studies (ds = 0.56 and 0.61, respectively), although there were no differences in initial self-efficacy between the latter two studies. These results suggest that students were initially least confident in their

ability to use immediacy as compared with challenges or interpretation, perhaps because immediacy felt less familiar prior to training.

In terms of final levels of self-efficacy, participants in the Chui et al. (2014) study were higher at post training than those in the Spangler et al. (2014) or Jackson et al. (2014) studies (ds = 0.62 and 0.45, respectively), with no differences between the latter two studies. These results suggest that students felt the most confident after training in their ability to use challenges, perhaps because challenges are easier to learn.

These results clearly show that undergraduate students increased in self-efficacy for using insight skills. We are not claiming that trainees were able to use the skills as effectively as experts would use them (we do not know of any such data), but rather that they had gained confidence in being able to use the skills therapeutically. Continued practice is undoubtedly necessary to refine the use of skills and use them with actual clients, and considerably more research is needed regarding the maintenance and refinement of skills. It is important to note that only a few researchers have examined maintenance of skills in general, and these studies are very old (Butler & Hansen, 1973; Collingwood, 1971; Gormally, Hill, Gulanick, & McGovern, 1975).

Other Outcomes of Training

In the Spangler et al. (2014) study on immediacy training, we found that students increased in perceptions of catharsis and group cohesion (ds = 0.27 and 1.00, respectively). Increases in catharsis make sense, despite the name of the construct, given that two of the items in the scale reflect increased immediacy, specifically, the presence and helpfulness of group members being able to speak openly with one another (e.g., "Learning how to share, in an honest and responsible way, how group members are coming across to me"). Increases in group cohesion also make sense given the results from psychotherapy research suggesting that processing the therapeutic relationship can result in changes in the client's relationships outside of therapy (see Hill & Knox, 2009). These results suggest that training in immediacy is effective not only for helping students increase in self-efficacy but also in terms of personal benefits.

In the Chui et al. (2014) study, students increased in rated quality of written challenges to written client stimuli ($d = 0.71$), whereas, as expected, they did not change in rated quality of written reflections of feelings to written client stimuli (relatedly, students changed in quality of reflections of feeling but not challenges following training in exploration skills). These data suggest that students learned how to write a good challenge during training. Although responding in writing to a written client stimulus is not comparable

with using a challenge to an actual client, these results provide evidence that students learned the mechanics of writing a high quality challenge.

Finally, in the Jackson et al. (2014) study, we created an analog counseling situation that students engaged in before and after training in interpretation. Classmates who served as clients were given a script and asked to role-play a client presenting a problem that involved material that could be interpreted in numerous ways. Helpers were instructed to give at least one interpretation during the final 5 min of the 20-min helping session. After the session, the client completed two items about interpretation taken from the Helping Skills Measure (Hill & Kellems, 2002). Analyses showed significant increases for training on both the client-rated ($d = 0.25$) and helper-rated interpretation items ($d = 0.40$). It is interesting to note that the effect size for client-rated interpretation is low compared with the effect sizes for the self-report measures; it is possible that because helpers were only able to offer one interpretation, clients did not have much opportunity to judge the helper's ability to use interpretations.

Correlations among changes in outcome measures were not significant in any of the three studies (changes in self-efficacy, changes in catharsis, and changes in group cohesion in the Spangler et al. [2014] study; changes in self-efficacy and changes in quality of written challenges in the Chui et al. [2014] study; and changes in self-efficacy and changes in helper- and client-rated interpretation use in the Jackson et al. [2014] study). Thus, clearly, these outcome measures are assessing different things. All seem to be conceptually valid measures of outcome, but they assess different aspects of the change process. Self-efficacy refers to confidence in using the skill and may reflect interest in, and motivation to, continue training, although such confidence may be unfounded. Cohesion and catharsis refer to facilitative aspects judged to be present in the lab group, although this may not translate to ability to use the insight skills. Quality of written challenges refers to the ability to write a grammatically correct skill, although this may not translate to ability to use the skill in a session. Lastly, helper- and client-rated interpretation use refers to ratings of how much helpers used interpretations in brief sessions with clients, which is influenced by how well the client was able to portray the scripted client and offer an opportunity for the helper to use the insight skill. Thus, it is apparent that all of the measures assess different aspects of the insight skills and, therefore, would not be expected to correlate highly.

Effectiveness of the Components of Training

In these three studies, we specifically included three of the four components of training (instruction, modeling, and practice) suggested by Bandura (1969, 1986, 1989, 1991) in his social cognitive theory. Because of the large class size and limited time, however, we did not systematically manipulate feedback

across the three studies. We note, however, that feedback was an integral part of the practice component and that students commented in their reflection papers about the helpfulness of the informal feedback attempts.

We assessed the effectiveness of the components through several methods. First, we examined changes in self-efficacy immediately after exposure to each of the components using quantitative methods in all three studies. Second, we assessed the effectiveness of all the components through qualitative analyses of reflection papers written after training in the Spangler et al. (2014) and Chui et al. (2014) studies. Third, we had retrospective ratings of the components in the Chui et al. (2014) and Jackson et al. (2014) studies. These retrospective ratings were important because students had experienced all of the components at this point and were able to judge them relative to each other. These relative retrospective ratings were above average for all components, suggesting that students valued all components. These ratings are not discussed further here because they did not discriminate well among components. We discuss the components in the relative overall order of effectiveness. Table 3 shows the results across the three studies.

Practice

From all the data collected, practice seemed to be the most effective component for helping trainees learn the target skill. Here, we discuss practice overall, and then the different types of practice.

Overall practice. In the qualitative analyses considering all types of practice together, 73% of the trainees in the Spangler et al. (2014) study indicated helpful aspects of practice, whereas 45% indicated unhelpful aspects of practice. Similarly, in Chui et al. (2014), approximately 45% of the students indicated both something they liked and something they disliked about practice.

Furthermore, in listing what they wanted more of in training, students in the Spangler et al. (2014) and Chui et al. (2014) studies often mentioned wanting more practice of all kinds. Students emphatically stated that although the other components were helpful, what really helped them learn the skill was practice. In fact, many students said that the skills seemed easy when they read about them, heard about them in the lecture, and observed them being used, but realized how difficult they were when they tried to implement the skills. One student wrote, "The only way to master them [the skills] is to practice them." Another wrote,

> I would focus less on the time spent in class explaining specifics and more on lab times spent practicing. I felt more confident after the lab than I did from just hearing the professor speak about how they [skills] may be used.

Table 3. Evaluations of Components Across Studies.

	Spangler et al. (2014)		Chui et al. (2014)	
	Most helpful (%)	Least helpful (%)	Liked (%)	Disliked (%)
Reading	11	8	93	50
Lecture	14	14	69	29
Video modeling	35	18	67	46
Practice	73	45	42-49	44
Feedback	19	1		
Support	24	27		

	Chui et al. (2014)		Jackson et al. (2014)	
Ratings of components	*M*	*SD*	*M*	*SD*
Reading	3.71	0.80	3.52	0.90
Lecture	4.02	0.70	3.84	0.80
Video modeling	3.74	0.87	3.50	0.97
Class practice	3.73	0.96	3.39	1.03
Lab group practice	4.46	0.70	4.45	0.74
Dyad practice	4.10	1.04	4.02	1.12
Feedback	4.13	0.79		

Note. The proportions in the top half of the table were calculated from qualitative judgments of narratives written after training. Ratings shown in the bottom half of the table were reported by students after training using a 5-point scale (1 = *strongly disagree*, 5 = *strongly agree*).

Practice at the end of the lecture class. The practice at the end of the lecture involved the class as a whole discussing what they liked and disliked about the course (Spangler et al., 2014), writing a challenge after listening to a "client" on a video and then having some students share their challenges with the class (Chui et al., 2014), and a fishbowl exercise in which the instructor first served as the helper for a volunteer client and then some students in the class served as helpers (Jackson et al., 2014). Effect sizes for these practice exercises for changes in self-efficacy were 0.34, 0.51, and 0.44, respectively, suggesting that all were helpful methods, but that writing the challenges might have been the most effective method.

Lab group practice. The practice in the entire lab group (about 8 to 10 students) involved either a positive feedback exercise (Spangler et al., 2014) or the lab leader playing the client and the students taking turns as helpers, facilitating exploration and later providing a challenge or interpretation (Chui et al., 2014; Jackson et al., 2014). In the Jackson et al. study, a student also served as a client for the lab group with other students offering exploration skills and then interpretations. This exercise comprised the whole 2-hr lab period in Spangler et al., and the first hour of the lab in the other two studies. The effect sizes for this exercise were 1.05, 0.36, and 0.50, respectively. The practice in the Spangler et al. study may have been more helpful because it was longer or because it was safer, given that it involved only positive or neutral feedback rather than having to challenge or interpret.

Dyad practice. Dyad practice was included in both the Chui et al. (2014) and Jackson et al. (2014) studies for the second hour of the lab group. In both studies, the "client" was given a script that involved a role that could potentially set the stage for the helper to give a challenge or interpretation. The helper was instructed to facilitate exploration for the first 15 min or so, and then use at least one of the target skill (challenge or interpretation) in the last 5 min. Roles were then reversed so that each student had the opportunity to be both a client and a helper. The effect sizes comparing changes in self-efficacy from the large lab group exercise to the dyad exercise were small, 0.24 and 0.37, respectively. The smaller effect sizes compared with earlier practice exercises could have been due to a ceiling effect in the ratings (ratings could not get much higher at this point). They also could have been due to difficulties experienced by clients attempting to play scripted roles and helpers having to use insight skills after such a short period listening to the client.

Instruction

Although reading and lecture both fit under Bandura's (1986) instruction component, we found very different results for them. Hence, we discuss the results for them separately.

Lecture. Lecture had large or medium effects on self-efficacy (0.83, 0.66, and 0.52, respectively, across the Spangler et al., 2014; Chui et al., 2014; and Jackson et al., 2014, studies). Lecture seemed particularly helpful for teaching immediacy, which we speculate was because immediacy was a more difficult construct to grasp than challenges or interpretation, and students may have needed to hear about it and ask questions to understand it.

In the qualitative data, 14% found lecture to be one of the most helpful components and 14% indicated that it was one of the least helpful components (Spangler et al., 2014), whereas 69% mentioned things they liked and 29% things they disliked (Chui et al., 2014). Some students also complained about the lectures not adding anything beyond the book. Many noted that even when they got something from the lectures, the lectures did not enable them to actually implement the skill.

Reading. Students were required to read the relevant chapter in the text prior to attending the 4 hr of in-class training. At the beginning of the lecture class, all students took a quiz, which they passed, indicating that they at least skimmed the chapter and learned the basic content (or that the content was obvious). The effect size analysis indicated minimal effects for reading in the three studies (0.01, 0.19, and 0.00, respectively). Furthermore, scores on the quizzes were not related to changes in self-efficacy in either the Chui et al. (2014) or Jackson et al. (2014) studies (it was not possible to calculate this correlation in the Spangler et al., 2014, study). In the Spangler et al. study, 11% of the students indicated that reading was one of the most helpful components, although 8% said that it was one of the least helpful. In the Chui et al. study, 93% liked some aspect of the reading, whereas 50% disliked some aspect.

Interestingly, from the data presented, it appears that reading may have been more effective for teaching challenges than for immediacy or interpretation. We speculate that students may have felt apprehensive about using challenges because they feared harming the therapeutic relationship, and that the reading helped them understand that they could challenge while supporting clients. In contrast, interpretations might have been something students already had a good grasp of, intellectually, prior to training. Furthermore, the readings seemed not to have been particularly helpful in explaining immediacy. According to one student, "Learning to do immediacy was difficult because unlike most of the other helping skills we have used, it is not something we use in our daily lives and is not a normal part of most conversations." Another student wrote, "I was having trouble learning the concepts. To me, it is a very abstract idea, and watching it is less helpful than doing it. It was also difficult to grasp why immediacy is helpful." For another,

> It simply seemed to be a difficult concept to explain. I understood what it meant to think and speak in the here and now, but I did not understand from the beginning exactly how therapists may use this technique in training.

Video Modeling

In the Spangler et al. (2014) study, we created our own videos because there were no existing videos of the four types of immediacy used in the Hill model. The "helpers" and "clients" were doctoral students who adhered loosely to scripts. In the Chui et al. (2014) study, we used clips of expert therapists from the video that accompanied the first helping skills text (Hill & O'Brien, 1999). For the Jackson et al. (2014) study, we used clips from the televised series, "In Treatment" (García, 2008), with each clip showing a therapist's use of one or more interpretations in a session with one client.

The effect size analysis of changes in self-efficacy indicated medium effects for the first two studies (0.42, 0.52) but no effects for the third (0.08). In the Spangler et al. (2014) study, 35% of the students cited modeling as one of the most helpful components whereas 18% said it was one of the least helpful. In the Chui et al. (2014) study, 67% liked some aspect of the modeling, and 46% disliked some aspect. In their comments, students in all three studies complained about the quality of the excerpts and criticized the therapists for not being perfect. In fact, the excerpts stimulated considerable discussion among the students in all three studies because they often saw things they did not like about how these therapists came across. In particular, students were quite caustic about the therapist in the televised series in the Jackson et al. study because they perceived that the therapist was disrespectful towards the clients and did not use helping skills according to the model the students were being taught. We think that it is a good thing that students were able to critically evaluate the therapist's use of skills.

Decrease in Self-Efficacy Between Lecture and Lab Classes

In all three studies, students decreased in self-efficacy for using the target skill between the end of the lecture class and the start of the lab (a period of 2 to 5 days), although they did not decrease to baseline levels. This decrease in self-efficacy between lecture and lab may have been due to the anticipation of discomfort or awkwardness in practicing advanced helping skills with classmates, particularly in the intimacy of a small-group setting. They may have felt confident after more passive exposure to the skill but felt more nervous when the time came for more intense and individual practice.

Summary of the Effectiveness of the Components of Training

In looking across all of the evidence, all of the components were effective in at least one of the analyses. For example, in the ratings completed at the end

of training, after students had been exposed to all the components and rated them at the same time, all components were rated as above average in terms of helpfulness (range from 3.71 to 4.46 on a 5-point scale). Putting all the data together, however, the most effective components appeared to be practice and lecture. Interestingly, there was a discrepancy among the students about the effectiveness of the components. For example, although reading was generally not rated highly, many students in the Chui et al. (2014) study found it to be very helpful. In contrast, although practice was generally considered helpful, it was not uniformly so for all students. These results suggest that there may be student variables that influence preference for different components of training.

Predictors of the Effects of Training

We have evidence about the influence of the ability of four variables to predict the outcomes of training from the correlations between predictor and outcome variables. We also had evidence about predictors from the qualitative findings.

Correlations Between Predictor Variables and Outcome Variables

We examined four predictors of the effects of training (final levels of outcome variables and change in outcome variables): initial levels of self-efficacy for the target skill (SE), prior helping experiences (PHE), attitudes toward learning helping skills (ALHS), and natural helping ability (NHM). PHE, ALHS, and NJM were intercorrelated modestly (.17 to .47); but SE was not highly correlated with the other three measures (.00 to .30). Thus, these measures were somewhat related but by no means measuring the same constructs.

Table 4 shows the correlations between the predictor and outcome measures across the three studies. The two predictors that replicated across at least two studies involved initial levels of self-efficacy for the target skill and prior helping experience. In Spangler et al. (2014) and Jackson et al. (2014), students with the highest initial self-efficacy for the target skill and with the most prior helping experience had the highest level of self-efficacy at post-training. Thus, those who started high in self-efficacy, and had more prior helping experiences, also ended high. In contrast, those with the lowest self-efficacy and highest prior helping experiences had the greatest gains in self-efficacy over the course of training. Thus, prior helping experiences seemed to enable students to profit more from training, perhaps because they were

Table 4. Correlations Between Predictor and Outcome Measures.

	Predictor measures			
	Initial SE	PHE	ALHS	NHM
Final levels of outcome measure				
Self-efficacy				
Spangler	.38**	.18*	.09	.15
Chui	.25	.16	.15	.30*
Jackson	.36**	.34*	.08	.10
Catharsis	.22**	.17	.03	.16
Cohesion	.16	.20*	.12	.22*
Quality of challenges	.13	.00	.10	−.00
Interpretation use				
Client-rated	.03	.08	−.09	.18
Helper-rated	.18	−.01	−.05	.10
Change in outcome measure				
Self-efficacy				
Spangler	−.75**	.05	.04	−.12
Chui	−.07	−.12	−.05	.12
Jackson	−.71**	.10	.05	−.21*
Catharsis	.16	−.03	.19*	−.11
Cohesion	.10	.00	−.12	−.03
Quality of challenges	.13	.07	.15	.01
Interpretation use				
Client-rated	.02	.04	−.12	.01
Helper-rated	−.17	−.11	−.01	−.13

Note. Cohesion and Catharsis were assessed in the Spangler et al. (2014) study, quality of challenges in the Chui et al. (2014) study, and interpretation use in the Jackson et al. (2014) study. Initial SE = self-efficacy measure completed before training on the target skill; PHE = prior helping skills measure; ALHS = attitudes toward learning helping skills measure; NHM = natural helper measure; Final level = measure completed at the end of training on the target skill; Change = measure completed after training minus self-efficacy measure completed before training.
*$p < .01$. **$p < .001$.

attuned to what it takes to be a good helper. Those with low self-efficacy gained more from training, perhaps because they had more room to grow, but those who began with more confidence also ended with more confidence.

It is not clear why initial self-efficacy and prior helping experience were not significant predictors in the Chui et al. (2014) study. Perhaps it was more difficult for students to estimate self-efficacy for challenges than for immediacy and interpretation.

Evidence From Qualitative Analyses

We got glimpses that a number of cultural variables may have influenced trainees' ability to learn insight skills. Students commented about how values in their families and cultures toward challenging or being immediate either facilitated or inhibited their ability to learn and use these skills. It appears that if families were open to challenging and being immediate with each other, students felt freer to use these skills, but if there had been a cultural proscription against the use of such skills, students felt very awkward and uncomfortable. Similarly, gender seemed relevant in that students who had been socialized to follow gender-appropriate expressions of how direct and open to be with others, had difficulty learning and using these skills. We do not want to make too much of these data at this point, given that data were not collected systematically. However, we suggest that future research on cultural variables is very much needed.

Limitations

Results of the three studies should be considered within the context of how they were conducted. The major limitations involve participants, measures, and procedures.

Participants

The researchers/instructors were all from the same counseling psychology doctoral program. All had previously taught helping skills courses and believed in the effectiveness of training, although none was trying to "prove" the benefits of training. Rather, they were more intrigued with figuring out how to improve training.

Instructors have different teaching styles, and even the same instructors often have distinct relationships with different classes because of class composition and how instructors relate to particular students and group dynamics. So, although instructors followed a standard syllabus, used the same materials presented in the same order, and all lab leaders underwent training on how to conduct the labs, differences in personality, experience level, and teaching ability of instructors and lab leaders may have affected outcomes of training. We did not evaluate the adherence or competence of the instructors in administering the training, although this would be an interesting direction for future researchers to pursue. In addition, all the instructors had been trained in helping skills by the professor and used the professor's book (the professor was also an instructor in two of the studies), so results may not generalize to other ways of

learning insight skills. Each of the studies included at least four instructors, which is far more than the sole instructor included in many of the early helping skills studies. Ideally, we would have nested time within students and students within instructors in our hierarchical linear models (HLM) so that we could have tested for instructor effects. However, with only four or five instructors for each study, three-level models were underpowered, so we constructed models without nesting students within instructors. We did test for differences among instructors by adding membership in a specific instructor's class as a predictor. Our results indicated that the instructor variable did not make a significant difference in model fit. Thus, for the purposes of these studies, the individual instructor was not a factor in change in self-efficacy during training.

The participants were undergraduate students at one large mid-Atlantic U.S. selective public university. The majority of students were female, European American, and graduating seniors (approximate age 22). Although all psychology majors in this department were required to take laboratory courses, they had a choice among labs (the helping skills course was a lab). Students thus self-elected to take the course, suggesting that natural helpers may have been more likely to enroll. Furthermore, random assignment to the course versus a waiting list was not possible, and students chose which instructor they wanted within the restrictions of space availability.

In addition, as all instructors know, class composition makes a difference in the cohesiveness of the class and the outcomes of learning. Especially in an experiential course such as helping skills, students learn from one another almost as much as the instructor, so having a sense of respect of fellow students is crucial.

Measures

In terms of measures, we relied primarily on the self-report measures of self-efficacy for the target skill as our outcome. As noted previously, however, the self-report of self-efficacy measures had limited correlations with other measures used in the studies. In addition, students completed the self-efficacy measure up to 10 times, so reactivity and demand characteristics may have influenced the data in that participants may have remembered their previous ratings and felt that it would be socially desirable to give higher rating as training progressed.

Procedures

We assessed changes for when there was no training (using a delay condition in Spangler et al., 2014, and Jackson et al., 2014, and using a baseline to prior to training assessment in Chui et al., 2014), but we did not include

no-treatment controls or comparison conditions, which would have provided a clearer estimate of the effects of treatment. In addition, the components were all presented within a consistent sequence (reading was followed by lecture, modeling, and practice), so we cannot determine order effects.

Clearly, conducting research within a naturalistic classroom setting is challenging. We do believe that the advantages of having large samples of more than 100 students and four or five instructors in each study offset some of the limitations already noted.

Implications

Implications for Training

One of our most exciting findings was the effectiveness of practice for teaching insight skills. Especially when retrospectively (at the end of training) rating the relative effectiveness of all the components, students in all three classes valued the opportunity to practice the skill. They felt that other components (reading, lecture, and modeling) were helpful in preparing them for using the skill, but it was through practice that they believed they really learned how to use the skills.

Practice was not included in the Hill and Lent (2006) meta-analyses of the components of training because not enough studies had included it, so this current set of studies provide new evidence for its effectiveness and provide justification for including it in training. Given the effectiveness of a variety of types of practice, one implication for trainers is to provide opportunities for different types of practice (responding to written vignettes, large group practice, dyad practice, and the use of real stimuli as well as scripted stimuli). Responding to written or video vignettes allows students an opportunity to practice and become comfortable without having to respond in the moment to another individual. Practicing in a small lab group offers trainees the opportunity not only to practice the skills in a safe environment but also to receive feedback and hear how others (lab mates, lab leader) use the skills. Practicing with classmates in dyads allows students a chance to practice in a realistic setting, especially if there is a safe environment. Having the opportunity to serve as volunteer clients enables trainees to experience what it is like to have the skills used on them and gain empathy for the client's experience in a helping session. Having different types of practice likely appeals to different types of learners.

All of the other components that were manipulated (reading, instruction, and modeling) were also helpful to some degree, so a second implication is to include all the components in training in hopes of reaching different types of

students. Of course, as indicated in some of the qualitative feedback, trainees had complaints about the specifics of the reading, lecture, and video clips. Trainers certainly should strive to provide the best possible stimuli, but we also note that it is never possible to please everyone with the stimuli. In fact, the ability to critically evaluate the stimuli is important in its own right. For students to look at a video clip and say what they did and did not like is important experiential information to them in terms of figuring out how they want to approach working with clients.

Although we did not systematically manipulate feedback, comments from the qualitative portions of the studies indicated that students valued the feedback given by instructors and lab leaders. These comments lead us to believe that instructors should provide feedback, although we can offer no evidence from this study about effective types of feedback.

Although the predictor variables were not consistent across all three studies, results from the Spangler et al. (2014) and Jackson et al. (2014) studies indicate that initial self-efficacy for a target skill and prior helping experience may provide trainees with more confidence in their skill post training, which instructors may want to consider in developing their training programs.

Finally, given that the effects of training seemed to vary across individuals, another implication is to tailor training to fit the audience. Rather than being rigid about how the skills "should" be used or taught, we encourage trainers to teach trainees to be personal scientists and determine what works for them. It is important for trainees to observe the effects of their interventions on clients, and on themselves when they are clients, to see what works.

Implications for Research

To obtain large enough samples, researchers might need to conduct multisite studies, which would be advantageous because it would improve the internal and external validity of the designs. Site differences and instructor differences could be studied to determine the potential effects of trainers and biases.

Clearly, researchers need to develop measures that are relevant for changes in insight skills. Because we cannot expect undergraduate students to be able to use these skills in sessions with real clients in the middle of psychotherapy (for whom insight skills are the most appropriate) given that they have not yet received enough training to ethically allow them to see clients, and because insight skills are typically not appropriate until after the exploration stage with real clients, new methods are needed. One approach would be to use trained actors serving as clients as they do in medical training (e.g., the trained actor goes in and presents symptoms, and the medical trainee has to

make a correct diagnosis in a short period of time). We could train actors to present scenarios that would be appropriate for helpers to use insight skills; the actor could judge the quality of the helper's use of insight skills, and trained judges could rate the quality of the helper's responses. Although such assessments have drawbacks (see Gormally & Hill, 1974), they could be designed to fit the needs of assessing the effects of training on specific skills.

In an effort to resolve the discrepancies between the self-efficacy ratings and the performance-based ratings in the Chui et al. (2014) and Jackson et al. (2014) studies, we could develop new measures that helpers and clients could complete that parallel the four-item self-efficacy measures. We would expect more congruent results if helpers and clients were responding to similar items for self-efficacy and performance. We also suggest that better measures need to be created to assess the effects of training.

Moreover, we need better measures of helper characteristics to predict who benefits from training. In these and past studies, researchers have used only self-report measures of helper characteristics, which is problematic because participants often lack self-awareness, and characteristics are often manifested differently across situations rather than being static traits. We suggest the use of more nonverbal and performance-based measures for future research. Future researchers could also explore other variables that might predict the effects of training, such as personal experience with counseling and psychotherapy. For example, real-life experience with receiving challenges in a therapeutic setting may improve one's ability to learn how to give them in an empathic manner.

Trainer (instructor) effects also need to be investigated. It could be that some instructors inspire students to learn the skills, whereas others are not effective for particular students (a trainer-student interaction). We might expect that instructors who are charismatic, attractive, empathic, and credible would have better effects than their counterparts, at least with some students. Of course, there may be interactions between instructor and student styles, given that different students like different types of instructors.

We also need to know more about the effects of training for action skills. Unlike insight skills, where students feel they have not had much prior experience, students typically come in to helping skills training thinking that they know how to give advice. So training may involve more of a tweaking of their skills rather than teaching skills that are relatively new.

Furthermore, other components of training need to be investigated. In particular, feedback needs to be tested. The role of coaching, peer support, and peer modeling, as noted in the qualitative feedback, would also be important to investigate. Another component of training that could be tested is the effects of transcribing and coding sessions for the helping skills used as well

as what skills they wish they had used, with instructor feedback about the accuracy of coding and suggested alternate responses. We use this exercise in both the undergraduate and graduate courses, and students anecdotally report it is helpful in allowing them to see new possibilities for what they could have done differently.

We did not investigate whether success in learning exploration skills influenced ability to learn insight skills. Nor did we examine how training in immediacy, challenge, and interpretation interacted with each other. These would be great ideas for future research.

Another fertile area for research is the role of trainee self-awareness in being able to learn the skills. In this line, we suggest that it would be helpful to test the effects of using interpersonal process recall (IPR) techniques (watching a video of a practice session and having the trainee talk about what she or he was thinking and feeling at the time and what he or she thinks the client was thinking and feeling at the time) during training. Trainees' awareness of their intentionality seems particularly important for the effective use of advanced helping skills.

In addition, it would be helpful to assess the effects of having or not having helping skills prior to other types of training. Although we assume that helping skills training is crucial, obviously some therapists learn how to become very good therapists without ever having helping skills training. Thus, comparing different paths to becoming therapists could be valuable.

Conclusion

We hope that other researchers will become excited about this area and begin programs of research. Given that helping skills training seems to be important for laying a foundation for future practice, we need to have a strong evidence base for its effectiveness. Although research in this area is daunting and would be easy to avoid, we encourage researchers to study training as it is typically practiced in the classroom for both undergraduate and graduate students.

Authors' Note

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

Declaration of Conflicting Interests

Funding

References

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

Butler, E. R., & Hansen, J. C. (1973). Facilitative training: Acquisition, retention, and modes of assessment. *Journal of Counseling Psychology*, *20*, 60-65.

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: Changes in self-efficacy and written challenges, helpfulness of training components, and predictors of the effects of training. *The Counseling Psychologist*, *42*, **XX-XX**.

Collingwood, T. R. (1971). Retention and retraining of interpersonal communication skills. *Journal of Clinical Psychology*, *27*, 294-296.

García, R. (Producer). (2008). *In treatment: Season one* [Television series]. Los Angeles, CA: Home Box Office.

Gormally, J., & Hill, C. E. (1974). Guidelines for research in Carkhuff's training model. *Journal of Counseling Psychology*, *21*, 539-547.

Gormally, J., Hill, C. E., Gulanick, N., & McGovern, T. (1975). The persistence of communication skills for undergraduate trainees. *Journal of Clinical Psychology*, *31*, 369-372.

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. (2009). Processing the therapeutic relationship. *Psychotherapy Research*, *19*, 13-29.

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., Spangler, P. T., Chui, H., & Jackson, J. (2014). Training undergraduate students to use insight skills (immediacy, challenges, interpretation): An overview of three studies. *The Counseling Psychologist*, *42*, **XX-XX**.

Jackson, J., Hill, C. E., Spangler, P. T., Ericson, S., Merson, E., Liu, J., . . .Reen, G. (2014). Training undergraduate students to use interpretation: Changes in self-efficacy and interpretation use, helpfulness of training components, and predictors of the effects of training. *The Counseling Psychologist*, *42*, **XX-XX**.

Spangler, P. T., Hill, C. E., Dunn, M. G., Hummel, A., Walden, T., Liu, J., . . .Salahuddin, N. (2014). Training undergraduate students to use immediacy: Changes in self-efficacy, helpfulness of training components, and predictors of the effects of training. *The Counseling Psychologist*, *42*, **XX-XX**.

Author Biography

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.

John 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.

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.