A New Paradigm for Psychotherapist Development: Alliance-Focused Training and

Facilitative Interpersonal Skills

A dissertation presented to

the faculty of

the College of Arts and Sciences of Ohio University

In partial fulfillment

of the requirements for the degree

Doctor of Philosophy

Matthew R. Perlman

August 2020

This dissertation titled

A New Paradigm for Psychotherapist Development: Alliance-Focused Training and

Facilitative Interpersonal Skills

by

MATHEW R. PERLMAN

has been approved for

the Department of Psychology

and the College of Arts and Sciences by

Timothy M. Anderson

Professor of Psychology

Joseph Shields

Interim Dean, College of Arts and Sciences

Abstract

PERLMAN, MATTHEW R., Ph.D., August 2020, Clinical Psychology

A New Paradigm for Psychotherapist Development: Alliance-Focused Training and Facilitative Interpersonal Skills

Director of Dissertation: Timothy M. Anderson

Research on current methods of therapist training have demonstrated little evidence that training enhances therapy outcomes. This study investigated the impact of a novel, research-informed program for therapist development that integrated elements of alliance-focused training (AFT) and facilitative interpersonal skills (FIS). The AFT/FIS training intervention incorporated didactics, discussion, role plays, and deliberate practice, in a flexible, therapeutic relationship-focused workshop format. Fifty-eight graduate-level therapy trainees and professional therapists from various helping fields were randomized to one of two brief trainings in a multi-site RCT: 1) the AFT/FIS workshop or 2) a training as usual (TAU) workshop. Participants were assessed on critical, relational therapeutic skills before and after the training. Analyses revealed that participants in the AFT/FIS training saw a marginally significant increase in overall therapeutic skills and a significant increase in specifically targeted post-training therapist skills (i.e. empathy, alliance bond capacity, and alliance rupture-repair responsiveness) compared to participants in TAU. Implications for future empirical investigations and training initiatives are discussed.

Dedication

For Dr. Jeremy Safran: a mentor and a friend.

Acknowledgments

I would like to thank Dr. Timothy Anderson for his constant guidance and support over the last five years. I have learned more than I thought possible about research and practice in clinical psychology thanks, in no small part, to his outstanding mentorship. I also need to thank my peers within the Psychotherapy and Interpersonal Process Lab: Kevin David, Caroline Gooch, and Shelby Martin for their hard work, their stellar coding, and above all, their friendship. I owe many thanks to Kate Foley, Scott Mimnaugh, Joshua Finkelstein, Joshua Maserow, and all my collaborators in the Safran Lab whose steadfast commitment, in the face of unimaginable adversity, was central to the success of this project. Their strength has been an inspiration to me. Finally, I would like to thank my friends, family, and partner for their boundless love and continual encouragement throughout this entire process.

Table of Contents

List of Tables

Page

List of Figures

Page

Overview

Growing clinical, research, and policy-making attention is being paid to the process of training therapists within helping fields, particularly clinical psychology; however, comparatively little empirical evidence has linked current training models to enhanced outcomes with therapy clients (Callahan & Watkins, 2018a). The focus of this study was to introduce and experimentally investigate a novel paradigm for therapist training. This paradigm was designed to enhance the interpersonal abilities of therapists through training and deliberate practice on the working relationship between client and clinician. By operating on research-informed constructs and methods (i.e. deliberate practice, therapist effects, and the working alliance), it is hoped that this innovative paradigm for clinical training may be more successful in building therapist expertise, and by extension, increase important therapeutic outcomes.

Efficacy of Current Training Models

Given contemporary trends in clinical psychology that point towards the critical importance of empirically-based assessment, it stands to reason that therapist training programs should be held to similar standards (Callahan & Watkins, 2018a). Of the limited investigations into the efficacy of standard clinical training practices, there is a decidedly mixed set of conclusions on improvement in therapeutic skills/outcomes (Callahan & Watkins, 2018b). Several recent studies have examined the issue of graduate training and therapy outcome with larger trainee samples and longitudinal research designs. Owen, Wampold, Kopta, Rousmaniere, and Miller (2016) conducted a naturalistic examination of 114 psychology trainees (including practicum-level, internship, and postdoctoral students) for an average of 45.31 months ($SD = 15.31$). The researchers found modest, yet statistically significant, improvement in outcome, according to client self-report outcome measures. Specifically, for each one year in training, clients reported a small, yet significant, increase in average outcome ($d = 0.04$). Unfortunately, this growth only applied to clients presenting with low levels of initial symptom severity. Trainees, across all developmental levels, did not improve in their ability to successfully treat high severity clients.

Goldberg and colleagues (2016a) conducted what may be considered the largest study-to-date on psychology training. The researchers tracked a large sample of 170 therapists, ranging in experience from practicum trainees to early career professionals, over a greater span of time ($M = 56.76$ months) than prior studies. Therapy clients completed the OQ-45 (Lambert & Finch, 1999), a widely used measure of clinical

distress, at each session. The researchers determined that therapists' outcomes with clients on the OQ-45 actually *decreased* by a modest yet statistically significant amount over time. The results also showed no differences in treatment dropout rates by differing levels of experience. Unlike the methodology of the study conducted by Owen and colleagues (2016), the procedure used by Goldberg and colleagues (2016a) was quasi-random, meaning that therapists and clients were, largely, randomly assigned – only client preferences for therapist gender were honored when possible. Initial severity, specific clinical presentation, and trainee experience did not factor into case assignment. The authors also controlled for the potential confounds of initial client severity, therapist age, therapist caseload, and outliers, yet the results remained significant that client-rated outcomes became marginally worse over time. These findings may be limited by the lack of control for other, unaccounted for factors such as adherence to treatments and theoretical orientation of the study therapists. This study is also limited by its examination of drop-out rates and client-reported clinical distress as the only outcomes.

Utilizing data from the Goldberg and colleagues (2016a) study, Erekson, Janis, Bailey, Cattani, and Pedersen (2017) examined a subset of 22 psychologists who provided services to approximately 4,000 clients as they progressed through different developmental stages of training (e.g. practicum, internship, etc.). The results of their multilevel analyses mirrored the findings of Goldberg and colleagues (2016a); therapists largely remained the same, or became less effective, as they progressed through various stages of their training.

      Ultimately, the limited available data fails to demonstrate robust, conclusive evidence in favor of current clinical training models (Callahan & Watkins, 2018b). It should be noted that research on clinical training is extremely limited by a lack of true experimental manipulation (i.e. few randomized clinical trials [RCTs]), lack of generalizable outcome measures, few replicated results, and a lack of true random assignment. More research is needed before firm conclusions can be drawn about current training practices in clinical psychology; however, professionals in other fields, such as medicine and aviation, do reliably show enhancements in key job-related outcomes over time (Ericsson, 2009). Reviewing literature on the development of professional expertise may hold significant promise for informing the training of therapists.

Developing Professional Expertise via Deliberate Practice

In a comprehensive program of research on professional development, Ericsson (2009) highlighted deliberate practice as a key mediator for the development of expertise across various fields. Miller, Hubble, and Chow (2017) provided a four-pronged definition of deliberate practice in psychology which included: 1) a targeted, systematic approach to improving performance, 2) a mentor or trainer to guide the process, 3) regular, immediate feedback on the development of targeted skills, and 4) repeated practice and refinement in skill utilization. Taken together, these four interrelated elements provide a scaffolding for the development of expertise.

Recent investigations of deliberate practice in psychotherapy have found that therapists who spend more time practicing therapeutic skills had better client outcomes than those who did not (Chow et al., 2015; Goldberg et al., 2016b). In their naturalistic examination client outcomes in a single Canadian mental health agency, Goldberg and colleagues (2016b) collected data from 153 therapists and over 5,100 adult outpatient clients who completed the OQ-45 during treatment. Throughout the seven-year data collection period, the agency implemented quality improvement strategies which included deliberate practice and feedback. Small increases in effects ($d = .03$) were found within the agency and the individual therapist caseloads over each year. Level of client severity, changes in staff, and therapist experience did not appear to impact the trends in increased outcome.

Given these preliminary findings, deliberate practice may serve as an important training mechanism for helping therapists to develop expertise. As the foundation of a

novel therapist training system, it stands to reason that deliberate practice would be most effective if trainees are instructed on relevant skills and capacities (e.g. those which are linked to clinical outcomes and have broad applicability across distinct types of therapy). Meta-analyses have shown that therapist effects (e.g. Baldwin & Imel, 2013) and the therapeutic alliance (e.g. Flückiger, Del Re, Wampold, & Horvath, 2018) are two constructs which contribute substantively to client outcomes across many forms of treatment. Research on therapist effects by Anderson and colleagues (2009; 2016a; 2016b) has found that specific, facilitative interpersonal skills (FIS) help to explain why some therapists routinely achieve better outcomes with their clients than others. Research by Muran, Safran, Eubanks, and Gorman (2018) has found that providing clinicians with alliance-focused training (AFT) improves treatment processes and outcomes. Therefore, education and deliberate practice on the relational capacities associated with FIS (Anderson et al., 2009; 2016a; 2016b) and AFT (Eubanks-Carter, Muran, & Safran, 2015; Safran & Muran, 2000; 2006) offer a theoretically and empirically-sound foundation for a novel therapist training program.

Alliance-Focused Training

The establishment and maintenance of the therapeutic alliance between client and therapist has long been described as an integral component in any successful psychotherapy (Bordin, 1979; Safran & Muran, 2000). A recent, large-scale meta-analysis of nearly 300 research reports found that the quality of the alliance accounted for approximately 8% of the variance in adult therapeutic outcomes (Flückiger, Del Re, Wampold, & Horvath, 2018).

Safran and Muran (2000; 2006) pioneered a relational theory and subsequent series of AFT techniques designed to help therapists to effectively identify and resolve moments of rupture, or misalignment/disharmony, in the therapeutic alliance. AFT draws on elements of mindfulness, interpersonal schemas, research on alliance negotiation, and empirically-supported interventions from relational psychodynamic and cognitive behavioral therapy. In practice, AFT attempts to help therapists build three alliance-focused capacities: self-awareness, affect regulation, and interpersonal sensitivity through training, supervision, and experiential exercises (Eubanks-Carter, Muran, & Safran, 2015). Self-awareness assists therapists in attuning to and understanding when a rupture is developing in a session. Affect regulation is critical for therapists to tolerate challenging emotional states, in both themselves and their clients, as ruptures unfold. By utilizing affect regulation techniques, therapists are better able to handle difficult moments of rupture with empathy and openness as opposed to hostility or conflict avoidance, either of which may significantly worsen ruptures. Finally, interpersonal sensitivity is tailored to help therapists discuss moment-by-moment changes with clients

during ruptures in a way that promotes increased understanding between client and

therapists, with a broader goal of generalizing to relationships outside of the therapy

dyad.

AFT incorporates elements from both relational/process and skill-focused training

protocols. This training modality intends to foster an attitudinal awareness of the therapy

relationship and facilitate positive therapeutic process. While the AFT approach

integrates and promotes specific therapist behaviors, the protocol does not prescribe

particular skills for use in specific situations. Instead, AFT encourages therapists to

engage in appropriate process-based interventions based on their unique, contextual

moment-by-moment understanding of any given interpersonal situation. AFT offers a

widely-adaptable, responsiveness-based approach to improving therapy process. This

model of training is in line with recent research evidence showing that personalization

(Norcross & Wampold, 2018) and flexibility (Boritz, Barnhart, Angus, & Constantino,

2017; Owen & Hilsenroth, 2014; Weisz et al., 2012) in treatment process correspond with

better treatment outcomes.

Muran, Safran, Eubanks, & Gorman (2018) conducted a multiple-baseline trial of

AFT in 40 therapist-client dyads engaged in a 30-session protocol of cognitive-behavioral

therapy (CBT). Therapists were third- and fourth-year graduate students in clinical

psychology who had already completed one year of training and attained fidelity in CBT

prior to the study. In this trial, clients were adults diagnosed with a range of cluster C

and PDNOS personality disorders. Outcomes were assessed throughout treatment using a

composite of client and therapist assessments on symptomatology, interpersonal

behaviors, and adaptive functioning. Results revealed that the introduction of AFT was followed by significant, medium to large improvements in key interpersonal behaviors within sessions by therapists (e.g. reduced criticism) and clients (e.g. reduced avoidance). Clients also demonstrated more expressiveness which shared significant interactions with session-level and end-of-treatment therapy outcomes. Given these results, it appears that AFT shows significant promise as a core element of an effective therapist training paradigm.

Facilitative Interpersonal Skills: Therapist Effects in Action

Meta-analyses have concluded that therapists account for, on average,

approximately 5-8% of the variance in clinical outcomes (Baldwin & Imel, 2013). This

variance is not explained by demographic-level predictors such as cumulative years of

clinical experience (Anderson et al., 2009; Wampold & Brown, 2005). Instead of

examining developmental levels and various subjective self-assessments, pioneering

work by Anderson and colleagues on facilitative interpersonal skills (FIS; 2009; 2016a;

2016b) has found that specific interpersonal characteristics of therapists predict clinical

outcomes.

FIS are a set of relational qualities that help to explain the therapist effect. Drawn

from clinical observation, theory, and applied research (Anderson, Ogles, & Weis, 1999;

Anderson & Strupp, 2015), the FIS can be defined through eight domains: verbal fluency,

hope/positive expectations, persuasiveness, expressed emotion,

warmth/acceptance/understanding, empathy, alliance bond capacity, and alliance rupture

repair responsiveness. Taken together, these eight interpersonal qualities assist therapists

to foster positive relationships with clients, regardless of theoretical orientation,

presenting clinical problem, or level of clinical training/experience.

FIS are measured through an innovative performance-based task where

participating therapists respond to a standardized set of pre-recorded stimulus video clips

of difficult interpersonal moments in therapy. Drawn from actual therapy sessions, and

re-enacted for research purposes, the stimulus clips are specifically designed to

encompass multiple types of interpersonal conflicts across dimensions of the

interpersonal circumplex (Leary, 1957) ranging from domineering/hostile presentations to submissive/overly-friendly situations. Participants are tasked with responding to each of these stimulus clips, and the recordings are systematically coded in accordance with the eight FIS domains to provide an average FIS score on a scale ranging from one to five. The FIS task is the first performance-based measure to successfully predict therapist effects (Anderson et al., 2009).

Anderson and colleagues (2009) studied a sample of 24 therapists and 1,091 clients at a university counseling center to examine the impact of FIS. The researchers assessed FIS of all study therapists. They also accounted for age (which generally correlates with experience), gender, percentage of time conducting therapy, self-reported social skills, and theoretical orientation. Utilizing hierarchical linear modeling, nesting sessions within clients within therapists, the results showed that when all the predictors were placed in a model together, FIS emerged as the *only* significant predictor of client outcomes, as measured via client self-reported distress on the OQ-45. This finding aligns itself with a general finding which points towards a lack of connection between demographic variables, particularly age/experience, and outcomes, even in samples which include professional therapists (Anderson et al., 2009; Wampold & Brown, 2005). This finding also suggests that FIS may be assessing capacities beyond basic social skills and instead may tap into a specific set of higher order abilities that correspond to positive therapeutic process and outcome.

Anderson and colleagues (2016a) incorporated FIS into a test of therapists versus paraprofessional helpers. The research team assessed 11 advanced (i.e. having at least

two years of training) graduate students in clinical psychology and 12 untrained (i.e.

having two years of graduate study in a non-helping field, such as chemistry or art

history) graduate students for their FIS. All the participating graduate students, both

from clinical psychology and non-helping fields, were selected for having either

significantly high or low FIS. Each of the graduate students worked with two clinical

analogs for seven sessions each. Clinical analogs were 65 undergraduate students who

were screened for having elevated levels of clinical distress, similar to what would be

expected in an outpatient clinic. In addition, 20 students served in a non-treated control

group. The results demonstrated that there was no difference between trained versus

untrained helpers in achieving positive therapeutic outcomes; however, helpers with high

FIS, regardless of training status, were significantly more likely to have positive

outcomes at the end of treatment and follow up. FIS, not clinical training status,

predicted client-reported alliances as well. Among many considerations, these results

highlight the potential for FIS to correspond with a unique set of abilities which remain

untargeted by current training programs.

Anderson and colleagues (2016b) utilized FIS to predict naturalistic therapy

outcomes. Forty-four new clinical psychology graduate students were assessed for FIS

within their first week of graduate training. The students began seeing clients over one

year after this initial assessment. From a sample of 117 clients, analyses revealed that

this initial FIS assessment could predict clinical outcomes over one year later. Clients of

high FIS student therapists demonstrated most marked benefit in shorter term therapies

with differences between high and low FIS therapists diminishing as therapy length

increased. The predictive power of therapist interpersonal behaviors has been supported

by a recent investigation by Schöttke, Flückiger, Goldberg, Eversmann, and Lange

(2016). The researchers were able to predict clinical outcomes *five years out* after

observer coding therapy-related interpersonal behaviors of postgraduate trainees in a

baseline assessment. These results suggest that therapist interpersonal skills have highly

impactful and enduring qualities in understanding therapist effects.

Utilizing the FIS Task in Training

Utilizing FIS in a training program could provide a link between expertise and

translatable therapeutic qualities. The FIS task allows new therapist training methods to

have standardization, which has been cited as a significant issue for developing therapy

expertise (Tracey et al., 2015). Given the significant range of possible outcomes and

client factors which impact therapy, it may be extremely difficult for clinicians to receive

meaningful feedback on their own expertise based on session work. As such, a

standardized, ecologically-valid training tool such as the FIS may offer a unique

opportunity for specific feedback, which Miller and colleagues (2017) have cited as a

main tenet of deliberate practice. FIS also has a compelling evidence base which

robustly links the tool, and the assessed therapist characteristics, to client outcomes

(Anderson et al., 2009; 2016a; 2016b). Having training that reflects meaningful clinical

outcomes has been noted as a crucial prerequisite for the development of therapist

expertise (Goodyear et al., 2017).

Integrating AFT and FIS

FIS, while explicitly an examination of therapist effects, are primarily based in a therapist's ability to navigate moments of interpersonal conflict or rupture within the client-therapist relationship. Both AFT and FIS place significant focus on therapist action in attending to the alliance. A combined paradigm, rooted in developing expertise in FIS while incorporating core elements of AFT and a deliberate practice framework, represents an optimal theoretical basis for training therapists given the strong interrelationship between the AFT and FIS paradigms and their well-established links to clinical outcomes.

By utilizing FIS task scores as a marker for outcome, an AFT/FIS training paradigm would have a reliable, well-validated, and robust predictor of clinical outcomes with which to evaluate the efficacy of training procedures. While some training methods have been shown to be effective at increasing use of specific target skills, these results have not consistently predicted improved client outcomes. Any positive results from a combined AFT/FIS training program would have clear, translatable implications for clinical practice.

While using client outcomes for training would be optimal for maximizing ecological validity, the difficulties of training with large numbers of clients needed for reliable assessments make it extraordinarily difficult to execute on practical grounds. The time gap between training knowledge acquisition and clinical use, need to identify meaningful therapeutic process markers, and systems required to reliably identify practice-based errors all contribute to the concerns with utilizing therapist-client

interactions as an assessment method for novel training protocols. Due to significant variability in client outcomes (e.g. Cuijpers et al., 2012), training studies would need to employ dozens of clients per therapist in order to adequately account for statistical "noise" in their attempt to locate the "signal" specifically generated by the effect of training interventions. Almost no studies of training which utilize clients as sources of outcome have been able to achieve this standard.

Even rigorous training studies have encountered limits to their interpretability due to their use of client outcomes as sources of measurement. In the Owen and colleagues (2016) training study, the authors were unable to specify how clients were assigned across clinical placement sites, meaning that clients and therapists may not have been randomly assigned. Additionally, the researchers noted that initial severity of cases slightly decreased over time, which may explain the perceived improvement in outcomes over time as opposed to any effect of training.

Given the constraints of utilizing client outcomes as an assessment method, the FIS task represents an excellent compromise between ecological validity and methodological control through standardization. The FIS clips have all been carefully drawn from actual therapy encounters and asks participants to actively take on the role of therapists, which provides a solid connection to actual clinical experiences. Further solidifying these connections are the number of high-quality research studies which have shown FIS to be a reliable, enduring predictor of clinical outcomes over traditional performance metrics, such as therapist age, training status, or years of clinical experience (Anderson et al., 2009; 2016a; 2016b).

Finally, an integrated AFT/FIS paradigm makes consistent use of several important theoretical constructs in the design and execution of training. The training is based around assisting therapists to navigate the widely-studied and robustly impactful construct of the therapeutic alliance (Flückiger et al., 2018). By employing regular elements of deliberate practice along with immediate, targeted feedback stemming from standardized FIS task clips, the integrated training aligns strongly with theoretical and empirical research on developing professional expertise (Ericsson, 2009; Goodyear et al., 2017). Grounded in theory, strong empirical research, and years of clinical practice, a successful therapist training program based on the synthesis of FIS and AFT could form the basis of a needed paradigm shift in clinical psychology training.

Research on AFT/FIS

In a recent, preliminary study of an AFT/FIS integration, a team of researchers at Ohio University and The New School for Social Research collected data on a sample of 42 undergraduate students in the New York City area who reported an interest in a career in psychotherapy or counseling (Perlman et al., 2019). Using an RCT design, the undergraduate students responded to four baseline FIS clips and were randomized into one of two training conditions: 1) a preliminary version of an AFT/FIS training protocol, or 2) a training demonstration video of an internationally recognized expert in cognitive therapy performing a session and discussing cognitive techniques/theories. After completing training, students responded to four post-training FIS clips that the students had not yet seen. Coders were pre-trained and reliable in FIS coding. Coders were also blinded to condition and pre-post training status of all clips. Analyses indicated that

students who completed the AFT/FIS training were significantly more likely to see an increase in post-training FIS than those in the expert therapy demonstration training condition. Participants saw particularly large gains in the FIS items of empathy, alliance bond capacity, and alliance rupture-repair responsiveness, even after controlling for familywise type I error. This initial study of undergraduates interested in clinical careers provided strong preliminary evidence about the potency of an integrated AFT/FIS paradigm. It was believed that a more robust and clinically-tailored AFT/FIS training provided to current graduate students and professional psychologists, counselors, and social workers would achieve similar, significant results over traditional training methods.

In addition to designing an effective integrated training, a study on AFT/FIS could also build on prior FIS-related training literature (e.g. Hill et al., 2016) to determine what other qualities of therapists may be related to trainability. Given that the AFT paradigm heavily incorporates elements of interpersonal complementarity into training, it may be that trainees' trait interpersonal styles would impact their performance. In their study of undergraduate trainees enrolled in a course on helping skills, Hill and colleagues (2016) found that FIS corresponded to self-efficacy in helping skill use while self-reported trait empathy corresponded to baseline helping skill use. As such, a secondary aim of this study is to examine potential relationships between self-efficacy, trait empathy, and trait interpersonal styles as they relate to FIS performance and AFT/FIS trainability.

Hypotheses

The main aim of this project was to empirically investigate the ability of an integrated AFT/FIS training program to improve FIS across a broad population of student/professional therapists. Several hypotheses were formulated based on the study protocol:

1. Participants in the AFT/FIS condition would achieve a significantly greater increase in their post-training observer-coded FIS scores over those in the control training. This will serve as the main hypothesis for the study.

2. Due to the alliance-focused nature of the AFT/FIS training, participants who completed the AFT/FIS would achieve the most significant increase in observer-coded FIS scores measuring "empathy", "alliance bond capacity", and "alliance rupture repair responsiveness" as compared with participants in training as usual (TAU).

3. Baseline observer-coded FIS level was hypothesized to moderate the relationship between training condition and post-training FIS scores, such that participants with higher baseline FIS would perform more strongly in AFT-FIS than participants with low baseline FIS.

4. Across both conditions, self-reported of FIS, self-reported interpersonal communion, self-reported interpersonal agency, and self-reported trait empathy scores would share a modest, but significant, correlation with baseline observer-coded FIS scores but would not share significant correlations with post-training FIS scores.

Methods

Participants

Fifty-eight participants were drawn from two populations of interest: graduate-level trainees ($n = 43$) and professional therapists ($n = 15$) who have completed graduate training in a social helping field (e.g. clinical psychology, social work, counseling, etc.). Participant ranged in age from 22 to 70 years ($M = 31.64$; $SD = 11.67$). Participants' self-estimated lifetime clinical hours ranged from 0 to 35,000 hours ($M = 2,229.64$; $SD = 5,363.28$). Four participants reported having no lifetime clinical experience at the time of their involvement in the study. Participants reported a wide range of primary theoretical orientations including cognitive/cognitive behavioral (37.9%), integrative/eclectic (27.6%), psychodynamic/psychoanalytic (17.2%), existential/phenomenological/humanistic (5.2%), interpersonal (1.7%), or other/none/still deciding (10.3%). Participants also reported their fields of work/study as clinical psychology (48.3%), social work (37.9%), counseling (10.3%), and counseling psychology (3.4%). Most of the sample (74.1%) was made up of student therapists while the remaining participants were professionals (25.9%).

A subset of participants (34.5%) had some level of prior experience with/exposure to the FIS task. Study facilitators were familiar with the levels at which participants were exposed to the task and coded each participant on a 4-point scale ranging from 0 ("no exposure") to 3 ("significant exposure"). Participants with "little" exposure had typically seen only a few recordings of a prior version of FIS clips several years before they enrolled in the study. Participants with "significant" exposure had typically seen and

responded to the current version of the FIS task within 12 months of participating and received feedback on their responses from Dr. Timothy Anderson. Five participants had "significant" levels of exposure. One participant reported limited, informal experience with AFT principles.

Participants were required to be at least 18 years of age and fluent in English, in order to successfully complete the FIS task. Participants were also required to be current students in, or graduates of, a graduate training program in a social helping field. Study investigators contacted potential participants via recruitment emails to nearby graduate training programs, flyers, word of mouth, and in-person advertisements in area clinics. To assist in diversifying characteristics of the sample, participants were recruited from two socioculturally and geographically distinct sites: Southeastern Ohio and New York City. Twenty-seven participants were recruited from Southeastern Ohio while 31 participants were recruited from New York City. Graduate-level trainees were given $50 for full participation in the study. Professional therapists were given the choice between two CE credits or $50 for full participation.

Measures

Demographics form. Participants were asked to report on age, gender, graduate or professional status, field of study, primary theoretical orientation, and approximate number of hours of lifetime clinical work.

Interpersonal Reactivity Index. (IRI; Davis, 1980). The IRI is a widely utilized self-report measure of empathy (Gerdes, Segal, & Lietz, 2010). The measure features 28 items rated by respondents on a 5-point Likert scale ranging from 0 ("Does not describe

me well") to 4 ("Does describe me very well"). IRI items may be summed to create a

total score ranging from 0 to 112. The IRI attempts to measure both cognitive (through

the "Perspective Taking" and "Fantasy" subscales) and emotional (through the "Empathic

Concern" and "Personal Distress" subscales) components of empathy. Psychometric

investigations report acceptable internal consistency (Cronbach's α ranging from .71 to

.77) with test-retest reliability ranging from .62 to .71 (Davis, 1980). The IRI subscales

demonstrated acceptable to good internal consistency in the current sample (Cronbach's α

ranging from .63 to .80).

Interpersonal Adjective Scale – Revised. (IAS-R; Wiggins, Trapnell, &

Phillips, 1988). The IAS-R is a widely used measure for the assessment of interpersonal

style. The IAS-R includes 64 adjectives (e.g. "distant" or "extraverted") with 8

adjectives corresponding to the octants of the interpersonal circumplex. Participants rate

their identification with each adjective on an 8-point Likert scale ranging from 1

("extremely inaccurate") to 8 ("extremely accurate"). Octant scores are created by

averaging the scores of relevant items. Octant scores can be used as part of a formula to

create two-factor interpersonal circumplex scores of "domineering" and "communion"

traits. Reliability estimates range from .73 to .86 for each of the octants (Wiggins, 1995).

The IAS-R has been described as an excellent fit to the interpersonal circumplex structure

(Gurtman & Pincus, 2000). Internal consistency for each octant ranged from acceptable

to good in the present sample (Cronbach's α ranging from .66 to .85).

Facilitative Interpersonal Skills – Self-report. (FIS-SR). This measure asks

participants to FIS clips to rate their own performance following participation in the FIS

task. The measure includes eight items, corresponding with each of the eight items of the

FIS observer rating system. Participants rate the extent to which they believe that their

responses were warm, emotionally engaging, hopeful, etc. on a 5-point Likert scale

ranging from 1 ("strongly disagree") to 5 ("strongly agree"). Item scores are averaged to

create an overall FIS-SR score. The FIS-SR has been utilized in one recent exploratory

study of FIS in undergraduate college students (Anderson, Perlman, McClintock, and

McCarrick, 2019). In that study, FIS-SR shared only a moderate correlation ($r = 0.45$)

with observer-coded FIS. In the present sample, internal consistency was good

(Cronbach's $\alpha = .83$).

Facilitative Interpersonal Skills – Observer. The FIS task is scored using

observational ratings of audio or audio-visual responses to various stimulus clips of

difficult simulated interpersonal clinical scenarios. Observers code participants'

responses along eight items: verbal fluency, emotional expression, persuasiveness,

warmth/positive regard, hopefulness, empathy, alliance bond capacity, and alliance-

rupture-repair responsiveness. The items are rated on a 5-point Likert scale ranging from

1 ("skill deficit") to 5 ("optimal presence of skill"). All ratings are initially set at 3 and

move up or down based on various features of the response being coded. Item scores are

averaged to create a mean FIS score. Internal consistency in the present sample was

excellent both before (Cronbach's $\alpha = .92$) and after (Cronbach's $\alpha = .91$) training.

For the present study, FIS was coded by two distinct teams of coders of trained

graduate students at Ohio University ($N = 4$) and The New School for Social Research (N

$= 2$). All coders were pre-trained at their respective sites over a several month period to

achieve sufficient reliability on codes (typically denoted as an ICC above 0.80 for the FIS

task). Actual cross-site consistency between coding sites demonstrated sufficient

agreement on an initial set of pre-collected clips coded by both teams (ICC = 0.91).

Within each sub-group, the Ohio University team (ICC = .81) and The New School team

(ICC = .82) each achieved sufficient reliability.

Coders met regularly during the data collection process to ensure continued

calibration of coding. A coding manual, which has been used in several other FIS studies

(Anderson et al., 2009; 2016a; 2016b) was also utilized in the training and coding

processes of the study to increase reliability of the coding. In the current sample, the two

coders at The New School achieved a final ICC of .61 on 26 participants coded. The four

coders at Ohio University achieved a final ICC of .83 on 31 participants coded. It should

be noted that ICCs are influenced by both the number of coders and number of ratings,

both of which were somewhat low in The New School coding group when compared with

prior FIS research (e.g., Anderson et al., 2009).

Post-training survey. Following both training conditions, participants completed

a brief survey which contained a 3-item multiple choice content retention check based on

critical aspects of the training they received. Participants needed to get at least two of the

three questions correct to have their data included in final analyses. The survey also

included several free-response items to gather feedback from participants about the

training. One item (labeled as "critique") asked "what challenges in the video clip

response task could be better addressed in future versions of the training?" A second

item (labeled as "general feedback") asked "what general feedback do you have

regarding the training your received?" For the "critique" item, any response that indicated dissatisfaction or provided new areas to include in future trainings was coded as a "negative" comment. All other training-related responses to this item were coded as "neutral." For the "general feedback" item, responses which solely voiced praise or appreciation for the delivery, content, format, or lessons provided through the training were coded as "positive." Responses which solely voiced criticism, dislike, or confusion at the delivery, content, format, or lessons provided through the training were coded as "negative." Responses which included both positive and negative feedback as defined above were coded as "mixed."

Procedures

Study procedures were approved by institutional review boards at both university training sites. Participants completed the project in groups of up to four. All participants completed an informed consent process prior to participating. No participants refused to participate or terminated their involvement early. Prior to being randomized into one of the two trainings, all participants completed the Demographics Form, Interpersonal Reactivity Index (IRI; Davis, 1980), and Interpersonal Adjective Scales – Revised (IAS-R; Wiggins, Trapnell, & Phillips, 1988).

Once completed, all participants were oriented to a baseline assessment of FIS based on a pre-randomized sub-set of four FIS stimulus clips. Participants completed the baseline FIS assessment in separate rooms without interruption. Sound screen machines were used to ensure that participants could not hear each other's responses. The baseline FIS task was run from computers placed in each room and recorded via a webcam

mounted on each computer. After all participants completed the baseline FIS responses, study administrators stopped the recordings. Research assistants later edited the clips into four distinct files (one per each of the participant's four separate responses). These clips were saved using naming conventions such that future coders did not know the training condition or whether the clip was recorded pre- or post-training. All clips recorded at each training site (Ohio University or The New School) were coded by teams at the other site to further mask participants' identities and blind coders to study conditions. Coders were also blinded to the pre or post-training status of all clips.

After completion of the baseline clips, participants completed the Facilitative Interpersonal Skills – Self-Report (FIS-SR) measure which asks participants to rate their own performance on the set of baseline FIS responses. Participant groups were then randomized to one of two conditions: an integrated AFT/FIS training program or an active control, training as usual (TAU), which demonstrated key aspects of evidence-based practice in conducting cognitive therapy for depression (e.g. Beck, Rush, Shaw, & Emery, 1979). Randomization occurred after initial measure administration to prevent bias in how study administrators interacted with the participants.

After completing either the AFT/FIS training or TAU, all participants completed a post-training assessment of FIS which operated in a similar fashion to the baseline FIS assessment (separate rooms, recorded via webcam, etc.). This second set of four FIS stimulus clips were new to the participants but matched to the baseline clip set in the type of interpersonal concern being addressed. This is to say that stimulus clips between pre- and post-training "corresponded" to each other to ensure a roughly equivalent

combination of interpersonal situations (e.g. dominant, submissive, hostile, agentic, etc. client concerns) for participants to respond to both before and after training. Finally, participants in both trainings completed a brief content retention check featuring three multiple choice questions related to core content from the training. A passing score on this check, denoted as two or more correct answers, ensured that participants were attentive during training.

The AFT/FIS training consisted of a 90-minute small group workshop. AFT/FIS training was led by four advanced graduate student therapists (teams of two provided training at each site) who had been trained in AFT/FIS in collaboration with the program's originators, Dr. Jeremy Safran and Dr. Tim Anderson. AFT/FIS trainers learned how to deliver the intervention through observation of AFT workshops along with discussion and practice of AFT and FIS core principles. The TAU workshop consisted of a 90-minute video of both a therapy demonstration and interview with an internationally recognized expert in cognitive therapy. Both conditions are described below.

AFT/FIS training. The AFT/FIS training included several interrelated components delivered in a 90-minute workshop. The training began with a brief lecture/didactic element which introduces key concepts such as the therapeutic alliance, alliance ruptures, meta-communication, and interpersonal sensitivity. This component of the training also featured critical thinking/discussion moments along with clinical examples to better elaborate on and illustrate the concepts. The AFT/FIS intervention also included in-depth video-based examples of alliance rupture repair conducted by Dr.

Jeremy Safran that had been widely utilized in other AFT workshops. Training facilitators paused the video at key junctures to discuss specific elements of the process, solicit feedback from participants on potential interventions, and encourage sharing of participant reactions. In keeping with the principles of deliberate practice, participants completed role plays of specific therapeutic skills from the training. Role play scenarios were adapted from an activity designed to train therapists to identify and manage therapeutic transference (Steinfeld & Safran, 2018) given that prior findings highlight the importance of skill practice in changing therapist behaviors (Bearman et al., 2010; Beidas & Kendall, 2010). Training facilitators provided specific, targeted feedback to participants as they conducted role plays to 1) reinforce and integrate knowledge of AFT principles and 2) encourage flexibility in navigating difficult clinical situations. Finally, the AFT/FIS training exposed participants to clips that they previously responded to (pre-training). This component was designed to help participants to reflect on and re-evaluate their performance in the context of newly learned AFT principles and allow for deliberate practice of re-navigating these difficult clinical situations.

Training as usual. The training as usual (TAU) workshop condition was matched for time (90 minutes) and given a more traditional training approach by having participants view a 50-minute tape of a highly experienced therapist deliver a full session of evidence-based treatment, cognitive therapy (CT). Following the demonstration of CT, study administrators led participants through a discussion about their reactions to the session. Afterwards, training groups watched a second portion of the video which featured the expert clinician participating in an interview about the underlying theory and

techniques demonstrated in the session. After the video, study administrators led a final

group discussion about ways to implement some of the observed evidence-based

techniques into their own practices.

Analytic Plan

An a priori power analysis for linear regression (α = .05; power = .80) was conducted using G*Power 3.1.2 (Faul, Erdfelder, Buchner, & Lang, 2009) to detect medium effects between the two training conditions. Medium effects were selected as a more conservative estimate of effect size based on the recent preliminary AFT/FIS investigation of undergraduate students (Perlman et al., 2019) to account for the potential influence of adding several significant covariates. In the preliminary study of undergraduates, effects for the AFT/FIS training over control training were large. The power analysis determined that a total sample of 55 participants would be required.

Data was analyzed between the training groups for post-intervention scores to determine if significant differences emerged in FIS. Baseline observer-coded FIS ratings were added as the sole variable in the first step of all regression analyses in order to better isolate the impact of training over pre-existing or "natural" FIS. A preliminary stepwise procedure tested for potential covariates to ensure that the final regression models included all relevant study variables. Tested covariates included: demographics (including age, gender, theoretical orientation, level of training, and level of clinical experience), FIS-SR, IAS-R, and IRI scores, prior FIS exposure, FIS stimulus clip order, and training site (Ohio University or The New School). Analyses were conducted on a final sample (n = 55) which excluded three participants from the full sample (N = 58): one participant did not pass the post-training content retention check and two participants were missing all four recordings of their baseline FIS responses. Four participants were missing a single recording from their four baseline or post-training responses. These

participants were included and had their FIS scores averaged using their remaining three video responses.

Results

Training Acceptability

Qualitative comments from study participants were coded to examine potential differences in critical and general comments regarding both trainings. In response to the "critique" item, a chi square analysis revealed no significant differences between the trainings in the amount of critical comments, X^2 (1, $N = 55$) = .006, $p = .940$. In response to the "general feedback" item, 17 of the 23 comments (73.9%) in the TAU group were positive, four (17.4%) were mixed, and two (8.7%) were negative. Nineteen of the 22 comments (86.4%) in the AFT/FIS group were positive, three were mixed (13.6%), and none were negative. A Fisher's exact test was used due to the multiple instances of small cell count. The test revealed no significant differences between the two trainings on positive, negative, or mixed comments, $p = .460$. Overall, both trainings appeared to be viewed by study participants as credible, acceptable, and judged as beneficial to therapy practice.

Training Effects on Overall FIS

A hierarchical linear regression analysis was conducted to examine the impact of the interventions, along with other potential covariates, on post-training FIS (see Table 1 for descriptive statistics). To guard against overfitting (Hawkins, 2004) a final model to the relatively small final sample ($n = 55$), 15 potential covariates were entered into a preliminary hierarchical linear regression model using a forward stepwise procedure to retain the most impactful/important of the potential covariates (see Table 2). The first variable entered into the preliminary model was baseline FIS. Baseline FIS was entered

first into all regression analyses in the present study, given that baseline FIS emerged as a highly significant predictor of post-training FIS (all p's < .001, variance accounted for ranging from 38-59%). Demographic variables, including age, gender, lifetime hours of clinical practice, theoretical orientation, training status, and field of study/practice, were entered into the second block. The third block of psychological/personality variables included level of prior exposure to FIS and self-reports of: FIS, interpersonal dominance, interpersonal communion, empathy. The fourth block consisted of study design variables and included the order in which FIS stimulus clips were seen by participants and the site at which training was conducted.

Gender was the only covariate that significant predictors of post-training FIS (identifying as male corresponded with significant increases in post-training FIS), after accounting for baseline FIS. This single covariate (i.e. gender) was entered into a hierarchical linear regression, after accounting for baseline FIS, with training condition (AFT/FIS versus TAU) to predict post-training FIS. The overall model containing baseline FIS, gender, and training condition significantly predicted FIS, $F(3, 51) = 16.68$, $p < .001$, $R^2 = .495$. Baseline FIS, $\beta = .64$, $t(51) = 6.46$, $p < .001$, and gender, $\beta = .26$, $t(51) = 2.57$, $p = .013$, significantly predicted post-training FIS, such that participants with higher baseline FIS and participants who identified as male had higher post-training FIS. Training condition emerged as a marginally significant predictor of post-training FIS, $\beta = .20$, $t(51) = 1.98$, $p = .053$, such that participants in the AFT/FIS condition trended towards having higher post-training FIS than participants in the TAU condition. No significant interactions were detected among the predictors. Standard diagnostic

procedures were conducted to detect outliers, influential cases, violations of test

assumptions, and multicollinearity. None of these diagnostic procedures indicated

violations of test assumptions.

Training Effects on Targeted FIS Items

For the second hypothesis, a multivariate analysis of covariance (MANCOVA)

was conducted to examine the impact of training condition and relevant covariates on the

eight rated items on the FIS task. Given that the AFT/FIS intervention was designed to

primarily impact a therapist's ability to understand their client's perspectives, form a

collaborative alliance, and repair ruptures, we wanted to examine if individuals in the

AFT/FIS training saw appreciable benefits to the corresponding FIS individual skills of

empathy, alliance bond capacity, and alliance rupture repair responsiveness. Baseline

FIS and gender were entered into the MANCOVA as covariates, given that these

variables had previously been determined to be significant predictors of post-training FIS.

Training condition was entered as the predictor variable. The omnibus test was

significant such that there was a significant difference detected between training

conditions in post-training FIS item-level scores, $F(8, 44) = 3.44$, $p = .004$, partial $\eta^2 =$

.384. On the subsequent item-level ANOVA tests, participants in the AFT/FIS training

saw greater increases on post-training empathy, $F(1, 51) = 10.55$, $p = .002$, partial $\eta^2 =$

.171, alliance bond capacity, $F(1, 51) = 11.89$, $p = .001$, partial $\eta^2 = .189$, and alliance

rupture repair responsiveness, $F(1, 51) = 10.71$, $p = .002$, partial $\eta^2 = .166$. These results

remain significant even after controlling for familywise type I error using the Bonferroni

procedure. No significant differences emerged between the training groups on any of the

remaining five FIS items (see Figure 1). Test diagnostics revealed that one FIS item,

expressed emotion, violated the assumption of equality of error variances. Even after

removing this item and re-running the analysis, all results remained highly similar

without any changes to statistical significance levels. No other test assumptions were

violated.

FIS Correlations

For the final hypothesis, a series of Pearson's bivariate correlations were run to

examine the relationships between bassline and post-training FIS and other variables of

interest (see Table 3 for a complete correlation matrix). As expected, based on prior

studies (e.g. Anderson et al., 2009; 2015), traditional predictors of experience, such as

age, lifetime hours of clinical work, and level of training were not significantly related to

either baseline or pos-training FIS. Baseline FIS was significantly related to post-training

FIS, $r = .626$, $p < .001$. FIS shared little relationship with other trait-like variables.

Baseline FIS was unrelated to self-reported interpersonal dominance, self-reported

interpersonal communion, and self-reported empathy. Interpersonal dominance shared a

small yet significant relationship to post-training FIS, $r = .287$, $p = .034$.

Discussion

Overall, we considered this initial test of the AFT/FIS intervention to be successful. The AFT/FIS intervention had a marginally significant effect on mean post-training FIS. Nonetheless, there were significant and robust increases on the specific facets of FIS that were targeted in the AFT/FIS training. The AFT/FIS training led to significant improvements in the precise skills which are most targeted in the training: empathy, alliance bond capacity, and alliance rupture repair. These item-level findings were highly consistent with prior research on the impacts of AFT/FIS (Perlman et al., 2019). AFT/FIS places a significant emphasis on intervening on relational concerns as they occur in the therapy space (i.e. alliance rupture repair responsiveness), which entails an increased understanding of one's own affective states and those of one's client (i.e. empathy) and an ability to work collaboratively to resolve the relational issues (i.e. alliance bond capacity). The results of the current study may be seen as the first training interventions to significantly improve FIS in therapist populations.

Unexpectedly, baseline FIS did not moderate the relationship between training condition and post-training FIS. This lack of moderation may be explained by the relatively low variability on FIS found in the present sample. It may also be the case that graduate and professional therapists, at any baseline skill level, were able to benefit from the AFT/FIS intervention, which speaks to the broad applicability of an AFT/FIS training model. In further support of this applicability, participants in the study varied widely on their level of training, theoretical orientation, and field of work/study, yet the AFT/FIS intervention was generally viewed positively by all who completed the training. This

paradigm may also be widely adaptable, at a theoretical level, to many other clinical training programs and modalities. Across almost all types of psychotherapy, the therapist and the therapeutic alliance are seen as critical mechanisms for change. As such, an AFT/FIS educational protocol could complement more technique-focused training for students and practitioners across major schools of psychotherapy (e.g. cognitive-behavioral, psychodynamic, client-centered, etc.). More research is needed to determine what influence baseline FIS and other interpersonal abilities have on therapeutic performance.

It is also interesting to note how many traditional therapist characteristics failed to predict pre- or post-training FIS, such as age, lifetime experience, theoretical orientation, level/type of clinical training, and self-report of FIS (i.e. self-efficacy, for a review, see Larson & Daniels, 1998). Self-report of FIS did not share a significant relationship to observer coded FIS, which is interesting given that the items are largely identical, only differing by the individual completing the rating and the type of measure (state or performance). Prior research has shown that student helpers (Anderson et al., 2019; Hill et al., 2016) and professional therapists (Hogue, Dauber, Lichvar, Bobek, & Henderson, 2014) may report their own use of specific interventions in a way that is highly inconsistent with those of observers. In the present sample, participants generally over-estimated their interpersonal skill given that observer-coded baseline FIS scores ($M =$ 3.18) were appreciably lower than self-reported FIS on those same stimuli ($M = 3.44$). These results point towards the continued need for observer coding to accurately determine FIS, as prior studies reveal that therapists typically over-estimate their clinical

abilities (Walfish, McAlister, O'Donnell, & Lambert, 2012). While commonly used for training and assessment in other fields (Rooney & Nyström, 2018), performance-based simulation tasks such as the FIS are relatively rare in clinical psychology. Personality variables including interpersonal styles and trait empathy also did not share robust relationships with FIS. Taken together, these results suggest that FIS, as a construct, may measure a unique capacity, outside of traditional demographics, level of training/experience, or personality factors which uniquely explains therapist performance in clinical contexts.

While extremely encouraging that FIS could be significantly enhanced through a brief (i.e. ~90 minute) workshop intervention, more research is needed to ensure that these training outcomes would be enduring and contribute to client outcomes. Prior research has shown that brief, workshop-based trainings typically have limited effects on clinician behavior and that long-term consultation is required to maintain learning (Beidas, Edmunds, Marcus, & Kendall, 2012). As such, this study should be considered a preliminary test to determine if FIS could be enhanced at all in therapist/trainee populations, given the trait-like nature of FIS as observed in prior research (e.g., Anderson et al., 2016a).

It is also important to place these results in the context of a larger series of FIS-based training interventions conducted over the last several years. In a recent preliminary investigation of using video-recorded models of high and low skilled demonstrations followed by repeated practice of FIS responses, Anderson and colleagues (2019) found modest to moderate effects for modeled practice over a no practice time-and-attention-

matched control in an RCT of psychology undergraduate students at a large, Midwestern university. To enhance training effects further, Perlman and colleagues (2019) developed and tested a preliminary version of AFT/FIS for an RCT involving a new sample of psychology undergraduates from a geographically distinct region. In this study, large effects were detected in favor of AFT/FIS over a demonstration of cognitive therapy for enhancing specific FIS. The current study builds off the prior two by targeting a more relevant training population (i.e. graduate trainees and professional therapists), involving a more robust control training, and utilizing a more comprehensive AFT/FIS intervention.

Gender emerged as a significant predictor of post-training FIS. This gender difference has not been found in any prior FIS research (Anderson et al., 2009; 2016a; 2016b). The result may best be considered an artifact of the present sample since only nine participants, 16.4% of the total sample, were male. With no more than five males in either training condition, a few, more extreme cases may explain the gender differences detected in the current study. While little research has linked therapist gender to treatment *outcomes*, there is some data to suggest that gender may influence the effect of therapist supervision/training related to implementation of specific evidence-based practices (Bearman et al., 2013). The fact that male identity predicted higher FIS is at odds with limited extant literature which tentatively points towards women, not men, as having greater abilities in key therapeutic capacities such as empathy (Rueckert & Naybar, 2008) and alliance-building (Dinger, Strack, Leichsenring, Wilmers, & Schauenburg, 2008). More research is needed to determine what relationship that gender has with training, FIS, and therapist effects more broadly.

These results also point towards the potential need for further, modular trainings designed to focus more explicitly on the other five FIS in order to see enhancement in those capacities. Based on the results of this study, further trainings are being developed as part of a larger Facilitative Interpersonal Relationship Skills Training system (FIRST; Anderson & Perlman, in press). The FIRST system is designed to incorporate foundational empirical knowledge, practical skills, expert demonstrations, role plays, and guided reflection on past FIS clip responses for all eight of the domains rated on the FIS performance task. The FIRST system entails a series of four, two-hour workshops, each targeting two particularly interrelated FIS items (e.g. warmth/acceptance/understanding and empathy), along with a series of complementary homework assignments to encourage therapists to translate learning from the trainings into their clinical work. Deliberate practice is infused throughout the trainings and homework via reflection activities and role plays on key relational capacities. The FIRST system is designed to be implemented over a multi-week period, with regular supervision and individualized deliberate practice plans in support of key interpersonal skill development in order to help ensure that any enhancements in FIS are long-term effects. Additional components, including training modules and novel FIS stimulus clips on multicultural concerns, are also in development as key facets of the FIRST system. Further need for more broad-based training paradigms can be seen in the empirical literature on training clinicians to navigate the therapeutic relationship. A recent meta-analysis of alliance-focused trainings/supervisions failed to find a significant effect for relational trainings impacting client outcomes (Eubanks, Muran, & Safran, 2018). It is hoped that a more

comprehensive, common factors-based system such as FIRST will be able to produce a

more robust impact on therapeutic process and outcome.

Limitations

The current study was limited in its small sample size, thus preventing the inclusion of all relevant covariates from the final regression model. While this study included graduate trainees and professional therapists from two distinct regions, a larger, more diverse sample would have increased our power to detect significant effects as well as improved the ecological validity of our results. Ecological validity may have also been impacted by the fact that a subset of participants had prior knowledge of FIS and exposure to older versions of FIS stimulus clips. We believe that the impact of this existing knowledge of FIS was relatively minor as statistical analyses did not support a significant relationship between prior exposure and observer-coded FIS scores pre- or post-training. Still, future research on AFT/FIS or FIRST systems should attempt to utilize FIS-naïve populations.

The study may also be critiqued on the basis of "teaching to the test" in the AFT/FIS intervention by providing participants with foundational knowledge in FIS domains including empathy and alliance capacities (both bond/collaboration-building and rupture-repair responsiveness). We believe that the results of this study are still important given the distinct differences between the use of the FIS task as an assessment tool and the AFT/FIS elements of training. AFT/FIS attempted to help participants foster key relational capacities for use in a wide variety of challenging situations. The FIS task assesses a therapist's ability to provide an open-ended response to unique, challenging clinical scenarios. Some of the concepts trained in AFT/FIS, such as meta-communicating with a client about an alliance rupture, would not be directly useful in the

FIS task. Further, if a participant simply repeated phrases or responses learned in the training in order to respond to the new stimuli of the post-training FIS assessment, it would likely have failed, given the critical need for personalized, situationally-tailored responses for each FIS clip. As such, it is more likely that genuine learning occurred as participants would have to have translated training materials into novel situations (i.e. entirely new FIS stimuli) to succeed on the post-training FIS assessment. This study's methodology also supports the idea that these results are genuine, given the fact that all FIS coders were completely blinded to the pre-post nature and training condition of all responses. Finally, given that FIS is robustly predictive of therapy outcomes (Anderson et al., 2009; 2016a; 2016b), it may be beneficial for clinicians to be trained to enhance relational abilities in accordance with those identified as FIS.

It should also be noted that the results of this study may have been impacted by a lower-than-expected intraclass correlation among a subset of the coding team. While the final ICC for coding conducted by the New York coding team was .61, which may be considered as "good" within the broader field of psychological measurement (Cicchetti, 1994), typical ICCs reported in prior FIS research typically strive for ICC values above .8. This lower ICC may, in turn, increase the variability or "noise" in part of the current study's FIS coding, thus potentially lowering the overall accuracy of the codes.

Additional limitations should be noted regarding the interpretability of the findings in this study. While FIS has been repeatedly shown to be a significant predictor of clinical outcomes, no study has yet linked AFT/FIS training to client outcomes directly. As such, we can only speculate as to whether therapists trained in AFT/FIS

would see better outcomes with their therapy clients. Future studies are being planned to test for potential impacts of FIS-based trainings on client outcome. It will also be important for some of these studies to take on a dismantling study approach to determine whether the format of the AFT/FIS workshop, the core content, or both are required to enhance FIS. In the current study, we were unable to pinpoint which component(s) drove the effects found in favor of AFT/FIS.

More research will also be needed to compare training interventions such as AFT/FIS or FIRST to more comprehensive alternative trainings. The TAU condition employed in the current study was, by no means, an all-inclusive comparison workshop nor a perfect model for all current clinical training methods. For example, while the TAU workshop included *some* elements of active learning, such as multiple points for group discussion, it did not include role play interventions while the AFT/FIS workshop did. Research has shown that active learning can be an important predictor of clinician behavior change after receiving training (Beidas & Kendall, 2010). Despite the limitations of the control condition, participants in TAU found the workshop to be helpful and did see their post-training FIS scores increase; however, prior research has shown that FIS scores tend to rise slightly over the course of task administration, even without training (Anderson et al., 2019). More research is needed to determine if elements of the TAU workshop hold benefits in enhancing FIS.

Ultimately, the results of this study provide preliminary evidence of potentially important targets and methods of clinical training. The AFT/FIS intervention shifts away from a focus on particular techniques or interventions and instead tries to enhance a

therapist's ability to apply flexibility in negotiating critical aspects of the client-therapist relationship. This training method appeared widely applicable and valued by a broad population of trainees and clinicians who came from different social helping fields and adhered to a wide array of theoretical orientations. By focusing on enhancing, not controlling, the critical and unique variability offered by therapists across key relational capacities, the AFT/FIS intervention may serve as the basis for a more robust clinical training paradigm, across distinct fields and treatment modalities, which may ultimately improve therapeutic process and outcomes.

Table 1

Descriptive Statistics for Variables (n = 55)

Variable	M	SD
1. Baseline FIS	3.18	.29
2. Age	31.45	11.77
3. Lifetime hours	1,946.71	5,101.62
4. Self-reported FIS	3.44	.56
5. Self-reported dominance	.07	1.97
6. Self-reported communion	.025	1.85
7. Self-reported empathy	71.84	9.68

	N	%
8. Prior exposure		
None	36	65.5
Little	7	12.7
Moderate	7	12.7
Significant	5	9.1
9. Gender		
Male	46	83.6
Female	9	16.4
10. Theoretical orientation		
Cognitive/CBT	22	40
Psychodynamic/analytic	10	18.2
Existential/humanistic	2	3.6
Interpersonal	0	0
Integrative/eclectic	15	27.3
Other/still deciding	6	10.9
11. Training status		
Student	42	76.4
Professional	13	23.6
12. Field of study/practice		
Clinical psychology	28	50.9
Social work	20	36.4
Counseling	6	10.9
Counseling psychology	1	1.8
13. Training site		
Ohio University	25	45.5
The New School	30	54.5
14. Clip order		
One	14	25.5
Two	17	30.9
Three	11	20
Four	13	23.6
15. Training condition		
TAU	30	54.5
AFT/FIS	25	45.5

Table 2

Summary of Forward Stepwise Regression Analysis for Selecting Covariates (n = 55)

Variable	β
Step 1	
[a]Baseline FIS	.64**
Step 2	
Age	-.17
[a]Gender	.25*
Lifetime hours of practice	.01
Theoretical orientation	-.04
Training status	-.11
Field of work/study	-.12
Step 3	
Prior exposure to FIS	.14
Self-reported FIS	-.11
Interpersonal dominance	.19
Interpersonal communion	-.06
Empathy	-.03
Step 4	
Training site	-.10
Clip order	.04

Note: all reported regression coefficients are from the final model of the forward stepwise procedure.

[a]denotes a predictor which was retained by the stepwise procedure.

*p < .05. **p < .01.

Table 3

Zero-Order Correlations for Variables (n = 55)

|	1	2	3	4	5	6	7	8	9	10	11
1. Post-training FIS	--										
2. Baseline FIS	.63**	--									
3. Gender[a]	.21	-.07	--								
4. Training condition[b]	.19	-.01	-.01	--							
5. Lifetime hours	.04	.04	-.06	-.16	--						
6. Training status[c]	.01	.18	.18	-.16	.54**	--					
7. Prior exposure[d]	.21	.09	.09	-.01	-.22	-.36**	--				
8. IAS-R dominance	.29*	.07	.25	.02	-.15	-.04	.04	--			
9. IAS-R communion	-.14	-.02	-.30*	.21	-.07	-.35**	.19	.03	--		
10. IRI empathy	.07	.22	-.16	.07	.22	-.14	-.00	-.21	.32*	--	
11. FIS self-report	.06	.23	.06	.12	.07	.20	-.29*	.23	.21	.09	--

[a]Gender was coded 0 = female; 1 = male.
[b]Training condition was coded 0 = TAU; 1 = AFT/FIS.
[c]Training status was coded 0 = student; 1 = professional.
[d]Prior exposure was coded as 0 = none; 1 = little; 2 = moderate; 3 = significant.
*$p < .05$. **$p < .01$.
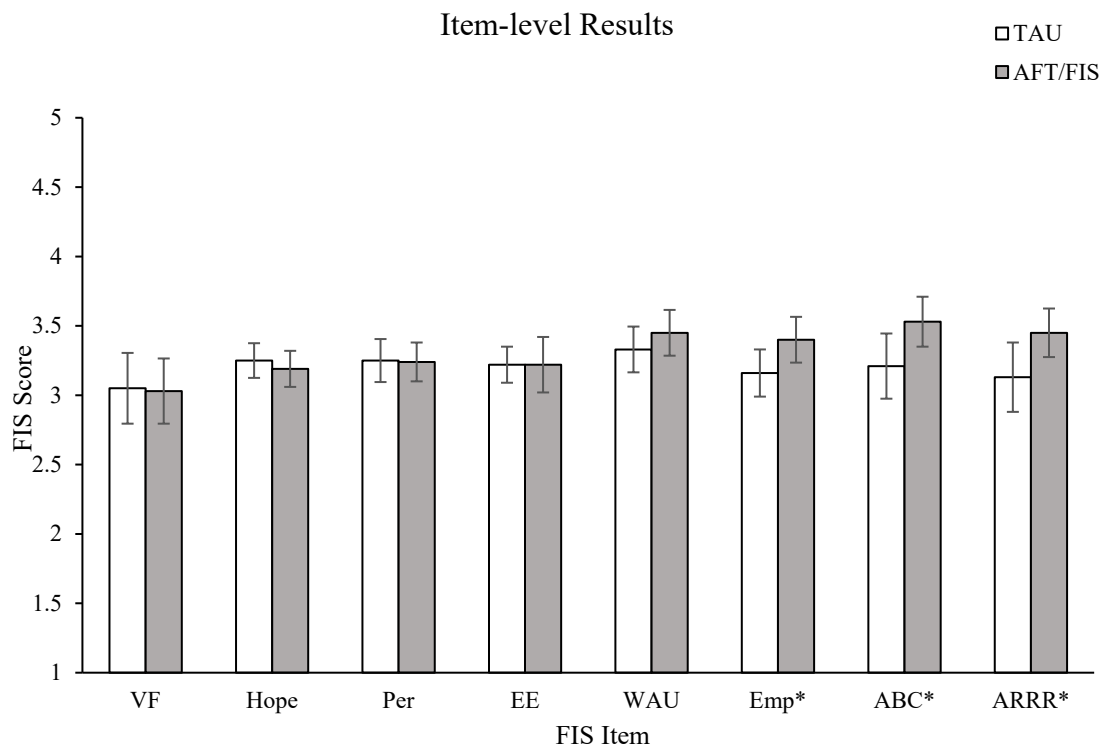
Figure 1. Item-level results for post-training FIS (*n* = 55).
*p < .05 (Bonferroni adjusted).

References

Anderson, T., Crowley, M. J., Himawan, L., Holmberg, J. K., & Uhlin, B. D. (2016a). Therapist facilitative interpersonal skills and training status: A randomized clinical trial on alliance and outcome. *Psychotherapy Research*, *26*, 511-529. doi: 10.1080/10503307.2015.1049671

Anderson, T., McClintock, A. S., Himawan, L., Song, X., & Patterson, C. L. (2016b). A prospective study of therapist facilitative interpersonal skills as a predictor of treatment outcome. *Journal of Consulting and Clinical Psychology*, *84*, 57-66. http://dx.doi.org/10.1037/ccp0000060

Anderson, T., Ogles, B. M., Patterson, C. L., Lambert, M. J., & Vermeersch, D. A. (2009). Therapist effects: Facilitative interpersonal skills as a predictor of therapist success. *Journal of Clinical Psychology, 65*, 755–768. doi: 10.1002/jclp.20583

Anderson, T., Ogles, B. M., & Weis, A. (1999). Creative use of interpersonal skills in building a therapeutic alliance. *Journal of Constructivist Psychology, 12*, 313–330.

Anderson, T., Perlman, M. R., McCarrick, S. M., & McClintock, A. S. (2019). *Modeling and deliberate practice of interpersonal skills improves facilitative interpersonal skills in short term responses*. Manuscript submitted for publication.

Anderson, T. & Perlman, M. R. (in press). Therapeutic interpersonal skills for facilitating the working alliance. In J. N. Fuertes (Ed.) *Working alliance skills for mental health professionals*.

Anderson, T., & Strupp, H. H. (2015). Training in time-limited dynamic psychotherapy: A systematic comparison of pre-and post-training cases treated by one therapist. *Psychotherapy Research*, *25*, 595-611. http://dx.doi.org/10.1080/10503307.2014.935517

Baldwin, S. A., & Imel, Z. E. (2013). Therapist effects: Findings and methods. In M. J. Lambert (Ed.), *Bergin and Garfield's handbook of psychotherapy and behavior change* (6th ed., pp. 258 –297). New York: Wiley.

Bearman, S. K., Weisz, J. R., Chorpita, B. F., Hoagwood, K., Ward, A., Ugueto, A. M., & Bernstein, A. (2013). More practice, less preach? The role of supervision processes and therapist characteristics in EBP implementation. *Administration and Policy in Mental Health and Mental Health Services Research, 40*, 518-529.

Beidas, R. S., Edmunds, J. M., Marcus, S. C., & Kendall, P. C. (2012). Training and consultation to promote implementation of an empirically supported treatment: A randomized trial. *Psychiatric Services*, *63*, 660-665.

Beidas, R. S., & Kendall, P. C. (2010). Training therapists in evidence-based practice: A critical review of studies from a systems-contextual perspective. *Clinical Psychology: Science and Practice*, *17*, 1-30.

Beck, A. T., Rush, A. J., Shaw, B. E, & Emery, G. (1979). *Cognitive therapy of depression*. New York: Guilford Press.

Bordin, E. S. (1979). The generalizability of the psychoanalytic concept of the working alliance. *Psychotherapy: Theory, Research & Practice, 16*, 252–260. http://dx.doi.org/10.1037/h0085885

Boritz, T., Barnhart, R., Angus, L., & Constantino, M. J. (2017). Narrative flexibility in brief psychotherapy for depression. *Psychotherapy Research*, *27*, 666-676.

Callahan, J. L., & Watkins Jr., C. E. (2018a). The science of training I: Admissions, curriculum, and research training. *Training and Education in Professional Psychology*, *12*, 219-230.

Callahan, J. L., & Watkins Jr., C. E. (2018b). The science of training II: Prepracticum and practicum training. *Training and Education in Professional Psychology*, *12*, 231-244.

Chow, D. L., Miller, S. D., Seidel, J. A., Kane, R. T., Thornton, J. A., & Andrews, W. P. (2015). The role of deliberate practice in the development of highly effective psychotherapists. *Psychotherapy*, *52*, 337-345. doi:10.1037/pst0000015

Cuijpers, P., Driessen, E., Hollon, S. D., van Oppen, P., Barth, J., & Andersson, G. (2012). The efficacy of non-directive supportive therapy for adult depression: a meta-analysis. *Clinical Psychology Review, 32*, 280-291. https://doi.org/10.1016/j.cpr.2012.01.003

Davis, M. H. (1980). A multidimensional approach to individual differences in empathy. JSAS *Catalog of Selected Documents in Psychology, 10*, 85.

Dinger, U., Strack, M., Leichsenring, F., Wilmers, F., & Schauenburg, H. (2008). Therapist effects on outcome and alliance in inpatient psychotherapy. *Journal of Clinical Psychology*, *64*, 344-354. doi: 10.1002/jclp.20443

Erekson, D. M., Janis, R., Bailey, R. J., Cattani, K., & Pedersen, T. R. (2017). A

    longitudinal investigation of the impact of psychotherapist training: Does training

    improve client outcomes?. *Journal of Counseling Psychology*, *64*, 514-524.

Ericsson, K. A. (2009). Enhancing the development of professional performance:

    Implications from the study of deliberate practice. In K. A. Ericsson (Ed.),

    Development of professional expertise: Toward measurement of expert

    performance and design of optimal learning environments (pp. 405-431). New

    York: Cambridge University Press.

Eubanks, C. F., Muran, J. C., & Safran, J. D. (2018). Alliance rupture repair: A meta-

    analysis. *Psychotherapy, 55*, 508-519. http://dx.doi.org/10.1037/pst0000185

Eubanks-Carter, C., Muran, J. C., & Safran, J. D. (2015). Alliance-focused training.

    Psychotherapy, 52, 169-173. http://dx.doi.org/10.1037/a0037596

Faul, F., Erdfelder, E., Buchner, A., & Lang, A. G. (2009). Statistical power analyses

    using G* Power 3.1: Tests for correlation and regression analyses. *Behavior*

    Research Methods, *41*, 1149-1160. doi: 10.3758/BRM.41.4.1149

Flückiger, C., Del Re, A. C., Wampold, B. E., & Horvath, A. O. (2018). The alliance in

    adult psychotherapy: A meta-analytic synthesis. *Psychotherapy, 55*, 316-340.

Gerdes, K. E., Segal, E. A., & Lietz, C. A. (2010). Conceptualising and measuring

    empathy. *British Journal of Social Work*, *40*, 2326-2343.

Goldberg, S. B., Rousmaniere, T., Miller, S. D., Whipple, J., Nielsen, S. L., Hoyt, W. T.,

    & Wampold, B. E. (2016a). Do psychotherapists improve with time and

experience? A longitudinal analysis of outcomes in a clinical setting. *Journal of Counseling Psychology*, *63*, 1-11. http://dx.doi.org/10.1037/cou0000131

Goldberg, S. B., Babins-Wagner, R., Rousmaniere, T., Berzins, S., Hoyt, W. T., Whipple, J. L., ... & Wampold, B. E. (2016b). Creating a climate for therapist improvement: A case study of an agency focused on outcomes and deliberate practice. *Psychotherapy*, *53*, 367-375.

Goodyear, R. K., Wampold, B. E., Tracey, T. J., & Lichtenberg, J. W. (2017). Psychotherapy expertise should mean superior outcomes and demonstrable improvement over time. *The Counseling Psychologist, 45*, 54-65.

Gurtman, M. B., & Pincus, A. L. (2000). Interpersonal Adjective Scales: Confirmation of circumplex structure from multiple perspectives. *Personality and Social Psychology Bulletin*, *26*, 374-384.

Hawkins, D. M. (2004). The problem of overfitting. *Journal of Chemical Information and Computer Sciences*, *44*, 1-12.

Hayes, A. F. (2012). PROCESS: A versatile computational tool for observed variable moderation, mediation, and conditional process modeling.

Hill, C. E., Anderson, T., Kline, K., McClintock, A. S., Cranston, S., McCarrick, S. M., ... & Gregor, M. (2016). Helping Skills Training for Undergraduate Students: Who Should We Select and Train?. *The Counseling Psychologist, 44*, 50-77.

Hogue, A., Dauber, S., Lichvar, E., Bobek, M., & Henderson, C. E. (2015). Validity of therapist self-report ratings of fidelity to evidence-based practices for adolescent behavior problems: Correspondence between therapists and

observers. *Administration and Policy in Mental Health and Mental Health Services Research*, *42*, 229-243.

Lambert, M. J., & Finch, A. E. (1999). The Outcome Questionnaire. In M. Maruish (Ed.), *The use of psychological testing for treatment planning* (pp. 831-869). Mahwah, NJ: Lawrence Erlbaum Associates Publishers.

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

Leary, T. (1957). *Interpersonal diagnosis of personality*. New York: Ronald Press.

Miller, S. D., Hubble, M. A., & Chow, D. L. (2017). Professional development: From oxymoron to reality. In T. Rousmaniere, R. K. Goodyear, S. D. Miller, & B. E. Wampold (Eds.), *The cycle of excellence: Using deliberate practice to improve supervision and training*. New York: John Wiley.

Norcross, J. C., & Wampold, B. E. (2018). A new therapy for each patient: Evidence-based relationships and responsiveness. *Journal of Clinical Psychology*, *74*, 1889-1906.

Owen, J., & Hilsenroth, M. J. (2014). Treatment adherence: The importance of therapist flexibility in relation to therapy outcomes. *Journal of Counseling Psychology*, *61*, 280-288.

Owen, J., Wampold, B. E., Kopta, M., Rousmaniere, T., & Miller, S. D. (2016). As good as it gets? Therapy outcomes of trainees over time. *Journal of Counseling Psychology, 63*, 12-19. http://dx.doi.org/10.1037/cou0000112

Perlman, M. R., Foley, V. K., Finkelstein, J., Anderson, T., Mimnaugh, S., David, K. C., Gooch, C. V., … & Safran, J. D. (2019). *Results from a preliminary test of AFT/FIS*. Manuscript in preparation.

Rooney, D., & Nyström, S. (2018). Simulation: A complex pedagogical space. *Australasian Journal of Educational Technology*, *34*, 53-64.

Safran, J. D., & Muran, J. C. (2000). *Negotiating the therapeutic alliance: A relational treatment guide*. New York: Guilford Press.

Safran, J. D., & Muran, J. C. (2006). Has the concept of the therapeutic alliance outlived its usefulness?. *Psychotherapy: Theory, Research, Practice, Training, 43*, 286-291. http://dx.doi.org/10.1037/0033-3204.43.3.286

Safran, J. D., Muran, J. C., & Eubanks-Carter, C. (2011). Repairing alliance ruptures. *Psychotherapy, 48*, 80 – 87. http://dx.doi.org/10.1037/a0022140

Schöttke, H., Flückiger, C., Goldberg, S. B., Eversmann, J., & Lange, J. (2016). Predicting psychotherapy outcome based on therapist interpersonal skills: A five-year longitudinal study of a therapist assessment protocol. *Psychotherapy Research, 27*, 642-652. doi:1.18/10503307.2015.1125546

Steinfeld, M., Safran J. D. (2018). *Teaching and experiencing Transference/ Countertransference*. Manuscript in preparation.

Tracey, T. J. G., Wampold, B. E., Goodyear, R. K., & Lichtenberg, J. W. (2015). Improving expertise in psychotherapy. *Psychotherapy Bulletin*, *50*, 7-13.

Walfish, S., McAlister, B., O'Donnell, P., & Lambert, M.J. (2012). An investigation of self-assessment bias in mental health providers. *Psychological Reports, 110*, 639–644. http://dx.doi.org/10.2466/02.07.17.PR0 .110.2.639-644

Wampold, B. E., & Brown, G. S. (2005). Estimating variability in outcomes attributable to therapists: A naturalistic study of outcomes in managed care. *Journal of Consulting and Clinical Psychology, 73*, 914-923. http://dx.doi.org/10.1037/0022-006X.73.5.914

Weisz, J. R., Chorpita, B. F., Palinkas, L. A., Schoenwald, S. K., Miranda, J., Bearman, S. K., ... & Gray, J. (2012). Testing standard and modular designs for psychotherapy treating depression, anxiety, and conduct problems in youth: A randomized effectiveness trial. *Archives of General Psychiatry, 69*, 274-282.

Wiggins, J. S. (1995). *Interpersonal Adjective Scales: Professional manual*. Odessa, FL: Psychological Assessment Resources.

Wiggins, J. S., Trapnell, P., & Phillips, N. (1988) Psychometric and geometric characteristics of the revised Interpersonal Adjective Scales (IAS-R). *Multivariate Behavioral Research, 23*, 517-530.

Appendix A: Measures

Demographics

Age: _____

Gender: _____

Training status:

[ ] Undergraduate student

[ ] Graduate student

      [ ] Master's degree program

[ ] Doctoral degree program

[ ] Professional mental health practitioner/therapist

Field of study/work:

[ ] Clinical psychology

[ ] Counseling

[ ] Social work

[ ] Other (please specify): _____

Primary theoretical orientation:

[ ] Psychoanalytic/Psychodynamic

[ ] Applied behavioral analysis/radical behavioral

[ ] Family systems/Systems

[ ] Existential/Phenomenological/Humanistic

[ ] Cognitive/Cognitive-behavioral

[ ] Integrative/eclectic

[ ] Other (please specify): _____

Years of clinical practice: _____

Approximate lifetime number of clinical intervention hours: _____

Interpersonal Reactivity Index

The following statements inquire about your thoughts and feelings in a variety of situations. For each item, indicate how well it describes you by choosing the appropriate letter on the scale at the top of the page: A, B, C, D, or E. When you have decided on your answer, fill in the letter on the answer sheet next to the item number. READ EACH ITEM CAREFULLY BEFORE RESPONDING. Answer as honestly as you can. Thank you.

ANSWER SCALE:

A	B	C	D	E
DOES NOT				DESCRIBES ME
DESCRIBE ME				VERY
WELL				WELL

Interpersonal Adjective Scales – Revised

On the front and back of this page are words used to describe people's personal characteristics. Please rate how accurately each word describes you as a person. Judge how accurately each word describes you on the following scale.

1	2	3	4	5	6	7	8
extremely	very	quite	slightly	slightly	quite	very	extremely
inaccurate	inaccurate	inaccurate	inaccurate	accurate	accurate	accurate	accurate

For example, consider the word "**BOLD**". How accurately does that word describe you as a person? If you think this is a *quite accurate* description of you, you would write the number "**6**" in the space next to the word BOLD.

6   BOLD

___

If you think this word is *slightly inaccurate* as a description of you, you would write the number "4". If it is *very inaccurate* you would write "2", and so on…

It is very important that you do not skip any.

1	2	3	4	5	6	7	8
extremely	very	quite	slightly	slightly	quite	very	extremely
inaccurate	inaccurate	Inaccurate	inaccurate	accurate	accurate	accurate	accurate

FIS – Self-Report

Instructions: Indicate how much each statement reflects your experiences with the responses that you gave to the clients in these videos. Select the number that best fits, from 1 (strongly disagree) to 5 (strongly agree).

In my responses to the videos, I…	**Strongly Disagree**				**Strongly Agree**
1. believe that my responses would have helped these clients	1	2	3	4	5
2. provided warmth and understanding	1	2	3	4	5
3. suggested new ways for the client to think about his/her problems	1	2	3	4	5
4. spoke in ways to let the client know that I wanted to collaborate with him/her	1	2	3	4	5
5. was hopeful, even when the client was negative or hostile toward me	1	2	3	4	5
6. felt confident that my responses to the client would be effective	1	2	3	4	5
7. felt comfortable when speaking and in a conversational tone	1	2	3	4	5
8. believe that I was able to speak with emotionally and in an engaging manner	1	2	3	4	5

Post-training Survey (TAU)

Please mark the correct answer to the following questions based on the training you received today.

1. Which of the following is a *primary* focus of the training you received?

    ○ a. Matching body language with your client

    ○ b. Understanding and addressing dysfunctional cognitive patterns in clients

    ○ c. Learning how to utilize your own life experiences as a method for helping clients

    ○ d. Working with individuals who are demanding

2. Which of the following is NOT a specific strategy that was taught in the training you received?

    ○ a. Highlighting alternative thoughts or explanations of issues

    ○ b. Examining evidence for or against particular cognitions

    ○ c. Trying to conceptualize core cognitive patterns/beliefs exhibited by clients

    ○ d. Utilizing silence as a method for reducing client anger

3. According to the training you completed today, which of the following is true about cognitive therapy?

    ○ a. Cognitive therapy works primarily by reducing emotional volatility in patients

    ○ b. Cognitive therapy is a treatment that can be utilized for a wide variety of psychological concerns

    ○ c. The outcome of cognitive therapy is largely determined by the client's assessment of the first session

    ○ d. Cognitive distortions may primarily be considered a manifestation of clients' anxiety about the future

Please complete the following free-response items based on your experiences in this training study.

4. In completing the therapy video clip response task, which elements of training (if any) did you draw on to help you respond?

5. What challenges in the video clip response task could be better addressed in future versions of the training?

6. What general feedback do you have regarding the training your received?

Post-training Survey (AFT/FIS Training)

Please mark the correct answer to the following questions based on the training you received today.

1. Which of the following is a *primary* focus of the training you received?

    ○ a.   Matching body language with your client

    ○ b.   Understanding and working through concerns within the therapeutic relationship

    ○ c.   Learning how to employ specific, evidence-based protocols at pre-set sessions

    ○ d.   Working with individuals who are demanding

2. Which of the following is NOT a specific strategy that was taught in the training you received?

    ○ a.   Explaining your rationales/methods of treatment when clients are unclear on the direction of therapy

    ○ b.   Utilizing your own internal experiences/reactions to help navigate therapeutic issues

    ○ c.   Communicating about the status of, or changes in, the therapeutic relationship

    ○ d.   Utilizing silence as a method for reducing client anger

3. According to the training you completed today, which of the following is true about the therapeutic alliance?

   ○ a.   The alliance is primarily formed because clients detect improvement over therapy

   ○ b.   The alliance can be thought of as a collaboration or negotiation between client and therapist about the directions, methods, and relationship of a treatment

   ○ c.   The alliance is largely determined by the outcome of the first session

   ○ d.   The alliance can best be considered a reflection or manifestation of the client's own internalized psychological concerns

Please complete the following free-response items based on your experiences in this training study.

4.. In completing the therapy video clip response task, which elements of training (if any) did you draw on to help you respond?

5. What challenges in the video clip response task could be better addressed in future versions of the training?

6. What general feedback do you have regarding the training your received?

Appendix B: Analyses

Alternate Test of Hypothesis One: Overall Test of Training Condition

In line with my original proposal, as an alternate test of hypothesis one, I completed a hierarchical linear regression which included all study predictors (see Table 1A). With so many predictors included, the sample size ($n = 55$) may have been insufficient for drawing meaningful conclusions and was at heightened risk for overfitting (Hawkins, 2004). As such, the forward stepwise procedure was utilized in the main paper to avoid these issues while still testing for potentially important covariates.

In this analysis, baseline FIS was entered as the first step. Baseline FIS was significant such that higher baseline FIS predicted higher post-training FIS. The second block consisted of demographic variables, including age, gender, lifetime hours of clinical practice, theoretical orientation, training status, and field of study/practice. None of these predictors emerged as significant after controlling for baseline FIS. The third block of psychological/personality variables included level of prior exposure to FIS and self-reports of: FIS, interpersonal dominance, interpersonal communion, empathy. None of these variables emerged as a significant predictor of post-training FIS. The fourth block consisted of study design variables and included the order in which FIS stimulus clips were seen by participants and the site at which training was conducted. Neither variable was significant. The final block, which added training condition, significantly predicted FIS, $F(1, 39) = 7.08$, $p = .011$, $R^2 = .067$. The results were such that participants in the AFT/FIS training saw significantly greater increases in post-training FIS than participants in TAU. After controlling for baseline FIS, demographics

(including training and clinical experience), psychological variables, prior FIS exposure, clip order, and training site, training site remained significant and accounted for approximately 7% of the variance in post-training FIS scores.

Table 1A

Summary of Hierarchical Linear Regression for Post-training FIS (n = 55)

Variable	β	ΔF	ΔR^2
Step 1		34.17**	.392
Baseline FIS	.678**		
Step 2		1.82	.115
Age	-.161		
Gender	.151		
Lifetime hours of practice	.275		
Theoretical orientation	-.036		
Training status	-.092		
Field of work/study	-.007		
Step 3		1.03	.054
Prior exposure to FIS	.148		
Self-reported FIS	-.095		
Interpersonal dominance	.213		
Interpersonal communion	-.164		
Empathy	-.019		
Step 4		.11	.002
Training site	.076		
Clip order	.107		
Step 5		7.08*	.067
Training condition	.282*		

Note: all reported regression coefficients are from the final model.
*p < .05. **p < .01.

Alternate Test of Hypothesis Two: Item-level Results

As an alternate test of hypothesis two (item-level analysis), all eight baseline FIS scores and gender were entered individually into the MANCOVA as covariates. In the main document, *mean-level* baseline FIS was utilized in order to align with the exact final model derived from analysis one. This choice also saved many degrees of freedom for the analysis. For both analyses, training condition was entered as the predictor variable. The omnibus test was significant such that there was a significant difference detected between training conditions in post-training FIS item-level scores ($F(8, 37) = 3.56$, $p = .004$, partial $\eta^2 = .435$). On the subsequent item-level ANOVA tests, participants in the AFT/FIS training saw greater increases on post-training empathy ($F(1,37) = 8.41$, $p = .006$, partial $\eta^2 = .161$), alliance bond capacity ($F(1, 37) = 9.29$, $p = .004$, partial $\eta^2 = .174$), and alliance rupture repair responsiveness ($F(1, 37) = 7.41$, $p = .009$, partial $\eta^2 = .144$). The results for empathy and alliance bond capacity remain significant even after controlling for familywise type I error using the Bonferroni procedure, while the result for alliance rupture repair responsiveness would be considered marginally significant. Utilizing the less conservative Holm procedure, all three results would remain significant. No significant differences emerged between the training groups on any of the remaining five FIS items. No test assumptions appear to have been violated.

Appendix C: Limitations

Removed Coder

As mentioned in the main document, one coder from The New was removed before including their data in the study analyses due to a lack of accuracy in coding. This decision was in-line with prior FIS studies where inaccurate coders are not utilized in data analysis; however, the ICC attained by The New School group slightly improves with the inclusion of this third coder (from .61 to .69). It should be noted that most FIS graduate student-level coding groups have consisted of four or more coders. ICC ratings are also influenced by the number of raters, not just the reliability of ratings. As such, The New School group, which only had three active members after unexpected early-study attrition, was particularly vulnerable to influence of even modestly correlated ratings. In other words, an inaccurate coder who still manages to vary their ratings in similar patterns to other coders may end up producing slightly higher ICCs like those observed in this case.

To test accuracy, I calculated a composite FIS mean and standard deviation score utilizing all published FIS coded data. Utilizing a sample of 310 ratings from five independent studies, I found a composite mean of 3.06 ($SD = .48$). The removed coder had a mean FIS rating of 4.04 ($SD = .44$). Based on this information, the removed coder's ratings were over two standard deviations higher than all other published FIS results. On these grounds, the coder's data was excluded from the dataset.

Removed True/False Item

As noted in the main document, the content retention checks were designed to feature a fourth true or false item for all study participants to complete. After examining the data, across both conditions, incorrect response rates were near 50%. The true or false items both featured confusing wording and asked about minute details related to the training, which may have led many participants to mark incorrect responses. As such, these items were dropped from the content retention check. The items are reproduced below:

True or false: a client who behaves in a way that's self-sabotaging would be an example of a cognitive distortion. [TAU]

○ True

○ False

True or false: a client speaking in a highly intellectualized manner to avoid facing difficult emotions would be an example of a confrontation rupture. [AFT/FIS]

○ True

○ False

Thesis and Dissertation Services