Journal of Consulting and Clinical Psychology
2016, Vol. 84, No. 1, 57–66

0022-006X/16/$12.00 http://dx.doi.org/10.1037/ccp0000060

A Prospective Study of Therapist Facilitative Interpersonal Skills as a Predictor of Treatment Outcome

Timothy Anderson, Andrew S. McClintock, Lina Himawan, and Xiaoxia Song
Ohio University

Candace L. Patterson
Psychiatric & Psychological Associates of Durham, Durham, North Carolina

Objective: This study examined whether therapists' facilitative interpersonal skills (FIS) would prospectively predict the outcomes of therapies that occurred more than one year later. ***Method:*** Therapists were 44 clinical psychology trainees who completed the FIS performance task and a self-reported measure of social skills in the initial weeks of their training. In the FIS task, prospective therapists were presented with a standard set of videos portraying clients in therapy. Verbal responses to these therapeutic simulations were recorded and then rated by trained coders. More than one year later, the therapists began providing psychotherapy to clients in a psychology clinic. Clients completed a symptom measure before each therapy session. ***Results:*** Using multilevel modeling, it was found that therapist FIS significantly predicted client symptom change. That is, higher FIS therapists were more effective than lower FIS therapists. However, subsequent analyses showed that this FIS effect was not uniform across all therapy durations; specifically, higher FIS therapists were more effective than lower FIS therapists over shorter durations (e.g., ≤8 sessions) but did not differ from lower FIS therapists in effectiveness for the small percentage of therapies that were longer-term (e.g., >16 sessions). ***Conclusions:*** Therapists' interpersonal characteristics may influence client progress in therapy.

What is the public health significance of this article?
This study found that therapist facilitative interpersonal skills (FIS), as measured by a performance task in the initial weeks of their graduate training, predicted the outcomes of therapies they delivered more than 1 year later. Although therapists higher in FIS were generally more effective than therapists lower in FIS, this effect was not apparent for a small percentage of therapies that were longer-term.

Keywords: common relational factors, therapist characteristics, therapist effects, psychotherapy outcome, working alliance

The psychotherapy literature has provided strong evidence that some therapists outperform others (Dinger, Strack, Leichsenring, Wilmers, & Schauenbirg, 2008; Kim, Wampold, & Bolt, 2006; Okiishi, Lambert, Nielsen, & Ogles, 2003; Wampold & Bolt, 2006). Overall, therapist effects account for between 5% to 9% of outcome variance (Baldwin & Imel, 2013; Kim et al., 2006). The clinical importance of therapist effects was recently illustrated by Baldwin and Imel (2013), who estimated that a typical client who sees one of the best 10% of therapists has twice the probability of recovery and half the probability of deterioration than if that client sees one of the worst 10% of therapists (see also Okiishi et al., 2003). Given the potential impact on public health, it is critical that we identify the specific therapist characteristics that account for these differential outcomes.

Unfortunately, the literature on therapist characteristics is sparse, with Castonguay (2013) recently describing it as "deplorable" (p. 54). That is, although client outcomes clearly vary by therapists, there is a lack of knowledge about the specific variables that are responsible for the differential effectiveness of therapists (Beutler, Machado, & Neufeldt, 1994). For example, therapists' age, gender, ethnicity, religion, marital status, clinical experience, and professional degree have not been consistently linked to client outcome (Beutler et al., 1994; Blatt et al., 1996; Huppert et al., 2001; Wampold & Brown, 2005). Because the source of therapist effects remains largely unknown, researchers (e.g., Baldwin & Imel, 2013; Wampold & Brown, 2005) have called for a renewed effort in isolating therapist characteristics that are associated with client outcome variability.

It may be useful to consider findings from the psychotherapy process-outcome literature as well as observations of high-performing therapists to identify those characteristics that account for therapist effects. For instance, an extensive body of research

This article was published Online First November 23, 2015.

Timothy Anderson, Andrew S. McClintock, Xiaoxia Song, and Lina Himawan, Department of Psychology, Ohio University; Candace L. Patterson, Psychiatric & Psychological Associates of Durham, Durham, North Carolina.

Correspondence concerning this article should be addressed to Timothy Anderson, Department of Psychology, Ohio University, Athens, OH 45701. E-mail: andersot@ohio.edu

has documented that certain processes in therapy, like empathy and alliance-building, are associated with client improvement (see Norcross, 2011; Horvath et al., 2011; Safran, Muran, & Eubanks-Carter, 2011). A recent meta-analysis (Elliott et al., 2011), involving a total of 3,599 clients, showed that about 9% of client outcome can be attributed to therapists' use of empathy. The therapists' expression of positive regard for the client also appears to be an important factor, as it accounts for about 7% of outcome variance (Farber & Doolin, 2011). It's plausible that the efficacious processes that occur in therapy (e.g., empathy, alliance-building) stem, in part, from the therapist's skill in creating such an environment. Moreover, therapists might differ in the degree to which they possess the interpersonal characteristics and skills that would give rise to these facilitative conditions. Thus, it stands to reason that some therapists outperform others because of pre-existing interpersonal characteristics that foster highly impactful therapy processes. Empirical evidence supports this reasoning, as the therapist's contribution to the alliance (i.e., what the therapist does with clients to create the alliance) is associated with better outcomes (Wampold & Imel, 2015). In other words, therapists who are able to form alliances across a range of clients also have better outcomes with clients, suggesting that therapists' interpersonal skills have an effect on therapy outcome.

Nevertheless, it is challenging to identify the interpersonal skills that might lead to outcome differences. In-session observation of therapist behavior seems problematic because the therapist's demonstration of relational skills could depend on the client seen as well as the therapist's relative ability to respond to different clients and situations (Bohart & Tallman, 2010; Boswell et al., 2013; DeRubeis, Brotman, & Gibbons, 2005; Imel, Baer, Martino, Ball, & Carroll, 2011; Stiles, 2013). Self-report measures of therapists' traits appear equally problematic. For example, therapists may perceive that reporting on their own skills, especially in their interpersonal abilities, could have personal and/or professional consequences (Stone et al., 2000). This possibility may introduce various biases (e.g., social desirability bias); some therapists might conceal perceived flaws or embarrassing information, whereas others may intentionally exaggerate their skills in a self-report (King & Bruner, 2000). In addition, many reports could be affected by blind spots and errors in memory and judgment (Tourangeau, 2009). The decision-making literature is rich in documenting these kinds of errors and biases among professionals, including psychologists (e.g., Loewenstein & Lerner, 2003). Because most research on therapist characteristics has relied on self-report, the conclusions that can be drawn from this body of research are limited. The present study will address this issue by using a prospectively administered, performance-based measure of therapist interpersonal skills.

Anderson et al. (2009) developed a performance-based measurement approach, referred to as FIS. For this task, which was informed by research on the appraisal of worker performance, a participant is asked to respond to a series of standard video clips that simulate challenging moments with therapy clients. The participant's responses to these clips are recorded and then rated by trained coders to determine the relative presence of eight facilitative skills (i.e., verbal fluency, emotional expression, persuasiveness, warmth/positive regard, hopefulness, empathy, alliance bond capacity, and alliance-rupture-repair responsiveness). This standard, performance-based procedure minimizes self-report bias and systematically controls client-related variability.

Empirical work to date has found that therapist FIS predicts alliance and outcome. Anderson and colleagues (2009) assessed the FIS of 25 therapists who treated 1,141 clients at a university counseling center. Using multilevel modeling of client-reported symptoms at each session, higher FIS therapists yielded greater rates of client improvement than lower FIS therapists. Therapists' self-reported social skills, age, gender, and theoretical orientation were unrelated to outcome. In a second study (Anderson et al., 2015), 23 therapists were selected for a) being high or low in FIS/social skills and b) for the presence or absence of clinical training (i.e., 2 + years of graduate study in clinical psychology vs. graduate study in other disciplines). Each therapist subsequently conducted individual therapy for seven sessions with two different clients. While the trained therapists were no more effective than untrained therapists, therapist FIS was a significant factor; clients who saw high FIS therapists had better outcomes (as evidenced by self-report and independent clinical assessments) at posttreatment and 3-month follow-up, as compared to clients who saw low FIS therapists. Moreover, high FIS therapists had higher client-rated alliances from the first therapy session, and unlike low FIS therapists, these alliance scores increased throughout the course of treatment. Collectively, these findings imply that therapists' preexisting relational skills may have an effect on the strength of the therapeutic alliance and treatment outcome.

One shortcoming of Anderson et al.'s (2009) study is that the FIS performance task was administered to therapists over a period of time that overlapped with the therapists' provision of therapy. Therefore, the predictive validity of FIS may have been due to transient, situational variables that were present during this period of time, rather than to therapist characteristics that were dispositional and stable. Evidence that FIS *prospectively* predicts client outcome, as found by Anderson et al. (2015), would provide better support for the notion that psychotherapy outcome is influenced by the therapist's preexisting skills. In addition, Anderson et al. (2015) selected independent groups of high and low interpersonally skilled therapists within an RCT design, however, selection of therapists used both performance-based data and self-report data, potentially limiting the conclusions that can be drawn from this study.

The current study addressed these limitations through the use of a prospective design where FIS performance-based data were collected from clinical psychology graduate students prior to their formal training and delivery of therapy. Furthermore, FIS was examined within a naturalistic setting of a training clinic, where therapists were beginning training and client diagnosis and treatments were not standardized. This choice was intentional because there is evidence that therapist effects are attenuated in tightly controlled, efficacy studies (Baldwin & Imel, 2013; Crits-Christoph et al., 2003; Wampold & Imel, 2015). In a meta-analysis of 46 studies, Baldwin and Imel (2013) found that therapist effects accounted for significantly more outcome variance in naturalistic, effectiveness studies (7%) than in efficacy trials (3%), perhaps suggesting that controls imposed on treatment delivery may have homogenized therapist performance. Accordingly, the present research used an externally valid research design; little to no restrictions were placed on the client sample, therapist sample, and session content (e.g., use of treatment approaches/techniques).

We hypothesized that in a prospective, naturalistic study, therapist FIS would predict client outcome, as measured by self-reported symptom change. We also examined additional therapist variables, such as self-reported social skills, age, gender, and theoretical orientation. Although speculation has existed about these therapist variables (e.g., Beutler et al., 1994), we did not anticipate that they would predict outcome and therefore made no hypotheses about them.

Method

Participants

Therapists. Prospective therapists were 50 graduate students enrolled in a clinical psychology Ph.D. program at a large Midwestern university. Six of these therapists (and their clients) were excluded from analyses because these therapists did not provide therapy to adult clients. The remaining 44 therapists were mostly (72.1%) female, had a mean age of 24.4 years (SD = 4.18), and identified as White/Caucasian (88.6%), Asian/Pacific Islander (6.8%), or Black/African American (4.6%). With regard to theoretical orientation, therapists self-identified as eclectic (38.2%), cognitive–behavioral (26.2%), humanistic (7.1%), psychodynamic (4.8%), or unidentified (2.4%). About 86% of therapists ultimately completed all phases of their clinical training (including clinical internship).

Clients. A total of 135 clients, who were either university students or members of the local community, were seen for psychotherapy by the therapists. Four clients were excluded from this study because they had been transferred form one therapist to another, thus potentially confounding the effects of two therapists. An additional 14 clients were excluded because they did not complete the symptom self-report measure during either of their first two psychotherapy sessions. The final sample of 117 clients was mostly (63.1%) female and had a mean age of 22.4 years (SD = 4.59, range: 18 to 42 years). Clients identified as White (88.9%), multiracial or other (5.9%), Hispanic/Latino (2.2%), African American (1.5%), Asian-Pacific Islander (0.7%), and Native American (0.7%). Clients sought treatment for a variety of reasons (e.g., anxiety, stress) and 58.7% reported a prior therapy experience in their lifetime. They attended a mean of 10.9 sessions (SD = 10.80, range: 1 to 60 sessions).

Client Measure

Outcome Questionnaire-45. The Outcome Questionnaire-45 (OQ-45; Lambert et al., 1994) is a 45-item general symptom measure that was completed by clients. Each item is rated on a 5-point Likert scale, ranging from 0 (*never*) to 4 (*almost always*). The sum of the items (after reverse coding selected items) forms the total OQ-45 score, which was used in the current study. The measure has demonstrated good internal consistency in prior studies (αs ranging from 0.70 to 0.93; Ogles, 1996), as well as in the current study (α = .96).

Therapist Measures

Social Skills Inventory. The Social Skills Inventory (SSI; Riggio, 1986) is a 90-item self-report questionnaire that measures social skills, including the ability to send and receive social messages through verbal and nonverbal channels. Items are scored using 5-point Likert scaling ranging from 1 (*not at all like me*) to 5 (*exactly like me*). The total of the items provides an overall indicator of social skills, which was used in this study. Test–retest reliability of SSI has been found to be high (rs ranging from 0.81 to 0.96) over a 2-week interval (Riggio, 1989). Convergent and discriminant validity for the SSI were supported in a series of studies conducted by Riggio (1986). In the present study, the SSI was completed by prospective therapists and had good internal consistency (α = .84).

FIS. FIS is an observational rating of audio responses provided by therapists to difficult simulated clients. There are eight items on the rating scale, all of which pertain to the therapist's skill in fostering facilitative conditions. These eight skill domains are verbal fluency, emotional expression, persuasiveness, warmth/positive regard, hopefulness, empathy, alliance bond capacity, and alliance-rupture-repair responsiveness. Each of these domains was rated on a 5-point Likert scale, ranging from 1 (*skill deficit*) to 5 (*optimal presence of the skill*). All ratings were initially anchored at 3 and were moved up or down the scale based on evidence of skills found in the audio responses. To increase reliability of these ratings, a coding manual was used that provided descriptions for each of the skills.

The FIS ratings were made by four coders, which included one doctoral-level researcher (Caucasian male), two graduate students (Chinese female and Caucasian male), and one undergraduate student (Caucasian female). Instruction in the FIS rating method occurred weekly over a 2-month period. Once there appeared to be sufficient agreement, ratings for the study commenced. The prospective therapist responses were rated in random order and in sets of 10. Each coder made their ratings separately and independently. Calibration meetings were held after each set of ratings, where discussion focused on those ratings that were most discrepant (i.e., typically greater than 1 point discrepancy). Final ratings for analysis in the study were a mean of all eight items, which were averaged across the four coders. Interrater reliability was acceptable for total FIS (intraclass correlation coefficient = 0.86), and the internal consistency of the eight FIS items was high (α = .96).

Procedure

Institutional review board (IRB) approval was obtained, and all ethical standards were followed. In total, data collection and analysis occurred over a 12-year period. Over the first 6 years, six cohorts of graduate students (see Therapists section) completed the Social Skills Inventory (SSI; Riggio, 1986) and the FIS performance task during their first two weeks of a clinical psychology PhD program. During the second year of training, therapists began providing psychotherapy to clients in a psychology department training clinic and continued to practice in the clinic through their third and sometimes fourth years. Thus, psychotherapy data (see Psychotherapy) were collected from the second to eighth year of this study. As per agreement with IRB, study data were not coded or analyzed until all participating therapists had completed their graduate course work or had left the program (i.e., approximately 12 years after the initiation of data collection).

FIS performance task. For the FIS performance task, prospective therapists responded to eight brief simulated therapy

situations, which had been developed in prior research (see Anderson et al., 2015; Anderson et al., 2009). The brief, standard FIS video segments were first drawn from a large archive of actual therapy sessions; these segments were selected because the interpersonal transactions between the client presentations were judged to be challenging for even seasoned therapists. Attempts were made to select a myriad of problematic interpersonal exchanges (e.g., hostile, controlling, submissive exchanges) that roughly represented the communication styles on the interpersonal circumplex.

Prospective therapists were first given a brief video role induction, in which they were instructed to act as if they were the therapist in the upcoming client video situations. Prospective therapists also were provided with brief background information for each of the simulated clients in the performance task and were then presented with the eight brief videos (two for each of the four simulated clients, each approximately 1-min in duration). Each simulated client was filmed from behind the left shoulder of a therapist in the lower right hand corner of the screen, and with the client directly facing the camera. After each of the videos, the image of the client's face was frozen with the superimposed words, "It's your turn to talk" appearing on the video screen. Prospective therapists then provided verbal responses, which were audio-recorded by a computer-generated program. Each prospective therapist completed the performance task in an empty, private room, although the prospective therapists were aware that their responses were being recorded. Attempts were made to create an environment that was ecologically similar to where therapy is conducted (e.g., often in the clinic therapy rooms where they would later practice). As noted, prospective therapists' responses were coded at a later date.

Psychotherapy. Approximately 13 months after the FIS measurement, therapists began working as trainees in the psychology department's training clinic. Assignment of clients to therapists was based on a number of factors (e.g., therapist availability, current caseload). All therapists received supervision, usually from clinical psychology faculty. For a typical caseload (e.g., three clients), therapists received approximately 1 hr of individual supervision and 2 hr of group supervision per week. The clinic operated under an eclectic treatment model in which specific treatment approaches were negotiated between supervisors and supervisees. Clients completed the OQ-45 immediately before each session. Length of treatment (M = 11.6 sessions) was highly variable (SD = 10.82) in this naturalistic setting.

Data Analysis

For OQ-45 scores, two different analyses were conducted. First, a two-level hierarchical linear model (HLM), with clients nested within therapists, assessed the improvement of clients as a function of the predictor variables. This analysis provided a more straightforward analysis of outcome by using only initial and termination OQ-45 scores.

Second, to model the change in the clients' OQ-45 scores over time, a three-level HLM with sessions nested within clients and clients nested within therapists was fitted to the data (Bryk & Raudenbush, 1992; Singer, 1998). An unconditional growth curve was fitted first to the data, and if there was a significant change over time, the effect of the predictor variables on the growth curve would be investigated. Because this study used a naturalistic design, the number of therapy sessions varied considerably. Baldwin et al. (2009) demonstrated that slopes of symptom reduction can be influenced by treatment length. Thus to account for the effect of treatment length on the growth curve, the total number of sessions attended by each client was included as a covariate.

To account for the different shapes that the growth curve might take, three different unconditional growth curves were fitted to the data: (a) a growth curve with a linear term of the session number only (i.e., a linear growth curve)—to assess the possibility that the growth decreases or increases in a constant rate over time, (b) a growth curve with both linear and quadratic terms of the session number (i.e., a quadratic growth curve)—to assess the possibility that the growth curve decreases first then increases or that the growth curve increases first then decreases over time; and (c) a growth curve with the log of session number—to assess the possibility that the growth curve decreases or increases at a faster rate during the first several sessions, then decreases or increases at a slower rate during the latter sessions. The fit of each model was assessed using information criteria (Akaike information criteria [AIC] and Bayesian information criteria [BIC]) to determine which model to use in subsequent analyses.

HLM analyses were also conducted for possible covariates on client demographic variables of age, gender, and prior therapy and "traditional therapist variables" that had been analyzed by Anderson et al. (2009) of age, gender, and theoretical orientation. We conducted a separate analysis for each of possible covariates entered at the appropriate level (i.e., client covariates were entered at client level and therapist covariates were entered at therapist level). Because the total number of completed sessions potentially affect the relationship between the OQ-45 and the covariates, the total number of completed sessions was also included as an independent variable in the three-level HLM. The nonsignificant covariates would be dropped for further analysis.

All HLM analyses were conducted with the PROC MIXED procedure of SAS Software, Version 9.3, of the SAS System for Windows using full maximum likelihood estimation procedures.

Results

Table 1 provides means, standard deviations, and zero-order correlations for the primary variables used in this study. To study the effects of the therapist variables on outcome, a two-level HLM was fitted to the data with the first session OQ-45 entered as a

Table 1
Descriptive Statistics and Zero-Order Correlations of Primary Variables

Measure	M	SD	Zero-order correlations			
OQ-45 First	76.86	16.52	1.00			
OQ-45 Last	64.82	14.62	.52***	1.00		
FIS	3.40	.52	−.00	−.20	1.00	
SSI	297.54	26.79	−.03	−.05	.33***	1.00
			OQ-45 First	OQ-45 Last	FIS	SSI

Note. OQ-45 First = Outcome Questionnaire-45 for first available session; OQ-45 Last = Outcome Questionnaire-45 for last session; FIS = Facilitative Interpersonal Skills; SSI = Social Skills Inventory.
*** $p < .001$.

level-1 covariate and therapist variables were entered on level 2. The model fitted was

$$\text{LastOQ}_{ij} = \beta_{00} + \beta_{01}(\text{Th})_j + \beta_{10}(\text{FirstOQ})_{ij} + [b_{0j} + e_{ij}],$$

where LastOQ_{ij} was the last session OQ-45 score for client i seeing therapist j; $(\text{FirstOQ})_{ij}$ was the first session OQ-45 score for client i seeing therapist j; $(\text{Th})_j$ was the therapist variable for therapist j (either FIS or SSI); β_{00} was the overall intercept (the average last session OQ-45 score); β_{01} was the coefficient of the therapist variable that indicated the number of unit increase/decrease in the last session OQ-45 score for each unit increase in the therapist variable; and β_{10} was the coefficient of the first session OQ-45 score that indicated the number of unit increase/decrease in the last session OQ-45 score for each unit increase in the first session OQ-45 score. The parameters inside the brackets were the random effects. In this model, the random effect accounted for the client variability around the overall intercept (b_{0j}).

For the FIS and SSI analyses, the random effects for the client variability around the overall intercept (b_{0j}) could not be estimated, and thus were not included in the model. There was a significant FIS effect as a predictor of OQ-45 outcomes in the model, but the effect for SSI was not significant (see Table 2).

For the three-level HLM, when fitting three different unconditional growth curves to the data, all three models showed a significant session number effect or a significant change over time. However, examination of the AIC and BIC showed that the growth curve with log session number provided the best fit among the three models tested ($\text{AIC}_{\text{linear}} = 9095.1$ and $\text{BIC}_{\text{linear}} = 9111.5$; $\text{AIC}_{\text{quadratic}} = 8970.5$ and $\text{BIC}_{\text{quadratic}} = 8997.8$; $\text{AIC}_{\text{log}} = 8962.5$ and $\text{BIC}_{\text{log}} = 8978.8$; a lower score indicates a better fit). Examination of the mean of OQ-45 scores by session (see Figure 1) provided further support of fitting a growth curve with log session number to the data. The figure showed that OQ-45 scores decreased at a faster rate in the early therapy sessions. Thus, the growth curve with log session was used in subsequent analyses. The final unconditional growth curve model was

$$(\text{OQ})_{tij} = \beta_{000} + \beta_{100}(\text{Log Session})_{tij} + [b_{0ij} + b_{1ij}(\text{Log Session})_{tij} + e_{tij}],$$

where $(\text{OQ})_{tij}$ was the OQ-45 score at time t for client i seeing therapist j; β_{000} was the overall intercept which was the average OQ-45 score at the beginning of therapy; β_{100} was the rate of change for the log session. The parameters inside the brackets were the random effects. In this model, the random effects accounted for the client variability around the overall intercept (b_{0ij}) and rate of change (b_{1ij}).

The random effects for the therapist variability around the overall intercept and rate of change could not be estimated, thus were not included in the model. In this model, the log session term was significant ($\beta_{100} = -5.30$, $SE = 0.64$, $F(1, 110) = -8.29$, $p < .001$).

After obtaining the best-fitting unconditional growth curve, the total number of session attended by each client was added as a Level 2 covariate and the therapist variable was added as a Level 3 predictor. Although the random effects part of this model was the same as the previous model, several terms were added to the fixed effects part of the model. The final model fitted was

$$\begin{aligned}(\text{OQ-45})_{tij} = {} & \beta_{000} + \beta_{001}(\text{Th})_j + \beta_{010}(\text{No of Sessions})_{ij} \\ & + \beta_{011}(\text{Th})_j(\text{No of Sessions})_{ij} \\ & + \beta_{100}(\text{Log Sessions})_{tij} \\ & + \beta_{101}(\text{Th})_j(\text{Log Sessions})_{tij} \\ & + \beta_{110}(\text{No of Sessions})_{ij}(\text{Log Sessions})_{tij} \\ & + \beta_{111}(\text{Th})_j(\text{No of Sessions})_{ij}(\text{Log Sessions})_{tij} \\ & + [b_{0ij} + b_{1ij}(\text{Log Session})_{tij} + e_{tij}],\end{aligned}$$

where β_{001} and β_{010} were the main effects for the therapist variable and number of sessions respectively. β_{011} was the interaction between the therapist variable and number of sessions. β_{101}, β_{110}, and β_{111} were the interaction of the therapist variable with the rate of change, the interaction of the number of sessions with the rate of change, and the interaction of the therapist variable with the number of sessions, respectively. β_{001}, β_{010}, and β_{011} tested the effects of the corresponding term to the overall intercept, which was the average OQ-45 score at the beginning of the therapy; whereas β_{101}, β_{110}, and β_{111} tested the effects of the corresponding term to the rate of change.

As akin to the two-level HLM analyses, client age, gender, and prior therapy and therapist age, gender, theoretical orientation, and SSI were not significant predictors in the three-level analysis and so were discarded from subsequent analyses. In terms of the influence of FIS on outcome, there was a significant FIS × Number of Sessions × Log Session interaction as well as an FIS × Log Session interaction (see Table 3).

The significant three-way interactions for Number of Session × FIS × Log Session suggest that the shape of the symptom reduction growth curve differed by FIS and therapy duration. To illustrate this, five different values were selected to represent five different therapy durations. We selected the same strata (i.e., 4-, 8-, 12-, 16-, and 20-session therapies) used by Baldwin et al. (2009). Of note, only 12.6% clients remained in therapy for 20 sessions.

Table 2
Result of the Two-Level Hierarchical Linear Model for OQ-45

Fixed effects		
	Coefficients (SE)	
Variable	FIS	SSI
Intercept (β_{00})	43.48 (9.549)***	34.65 (14.707)*
Predictor variable (β_{01})	−4.56 (2.151)*	−.02 (.045)
First session OQ-45 (β_{10})	.45 (.083)***	.46 (.072)***
Random effects		
	Variance estimates (SE)	
Term	FIS	SSI
Between patient		
Intercept (σ^2_{b0j})	—	4.70 (13.201)
Residual (σ^2_{eij})	139.76 (21.438)	149.20 (23.481)

Note. OQ-45 = Outcome Questionnaire-45; FIS = facilitative interpersonal skills; SSI = Social Skills Inventory.
* $p < .05$. *** $p < .001$.

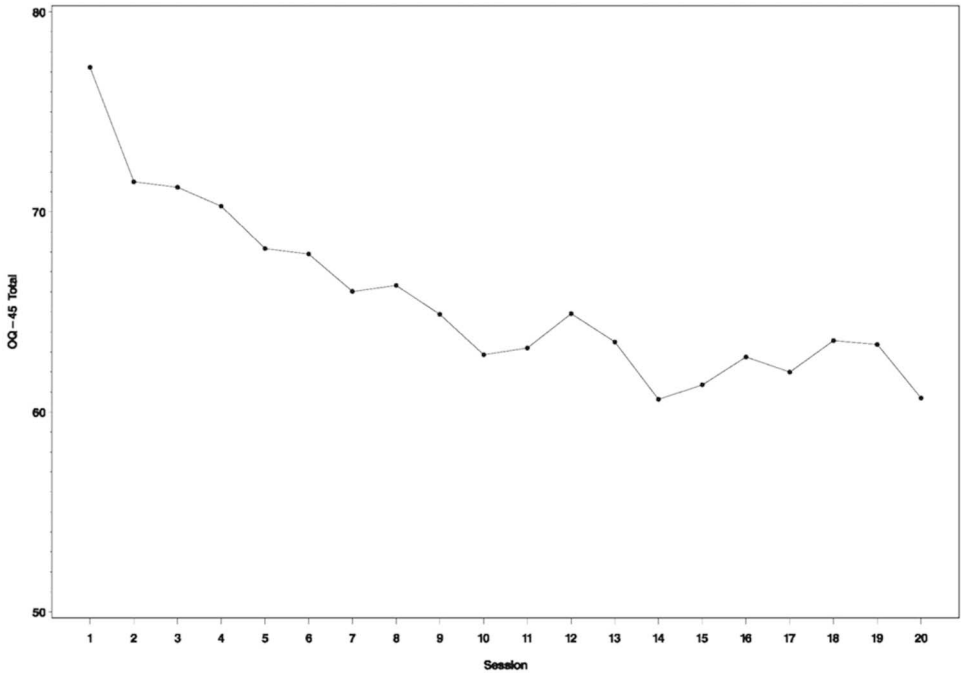

Figure 1. Mean Outcome Questionnaire-45 (OQ-45) scores by sessions.

To examine the effect of FIS on the shape of the growth curve, we used one standard deviation below FIS mean to represent low FIS, and one standard deviation above FIS mean to represent high FIS. Given that mean FIS = 3.401 (*SD* = 0.516), low FIS = 2.885 and high FIS = 3.937.

As depicted by Figure 2, the effect of FIS on client symptom reduction was not uniform across all therapy durations. Within the first three strata (i.e., 4-, 8-, and 12-session therapies), clients of high FIS therapists reported a substantially faster rate of improvement and better end-of-treatment outcomes than clients of low FIS

Table 3
Result of the Three-Level Hierarchical Linear Model gor OQ-45

Fixed effects		
	Coefficients (SE)	
Variable	FIS	SSI
Intercept (β_{000})	74.73 (15.748)**	92.08 (23.919)**
Predictor variable (β_{001})	.84 (4.786)	−.05 (.081)
No. of sessions (β_{010})	.48 (1.272)	−.74 (1.481)
Predictor Variable × No. of Sessions (β_{011})	−.18 (.378)	.00 (.005)
Log session (β_{100})	14.83 (7.093)*	−2.02 (10.493)
Predictor Variable × Log Session (β_{101})	−6.70 (2.174)**	−.01 (.036)
No. of Sessions × Log Session (β_{110})	−1.13 (.510)*	−.20 (.570)
Predictor Variable × No. of Sessions × Log Session (β_{111})	.37 (.153)*	.00 (.002)
Random effects		
	Variance estimates (SE)	
Term	FIS	SSI
Between patient		
Intercept (σ^2_{b0ij})	216.01 (36.432)	237.59 (34.818)
Log session (σ^2_{b1ij})	28.13 (5.944)	30.82 (5.803)
Covariance ($\sigma_{b0ijb1ij}$)	−40.82 (12.058)	−41.48 (11.234)
Residual (σ^2_{etij})	37.38 (1.804)	37.77 (1.630)

Note. OQ-45 = Outcome Questionnaire-45; FIS = facilitative interpersonal skills; SSI = Social Skills Inventory.
* $p < .05$. ** $p < .01$.

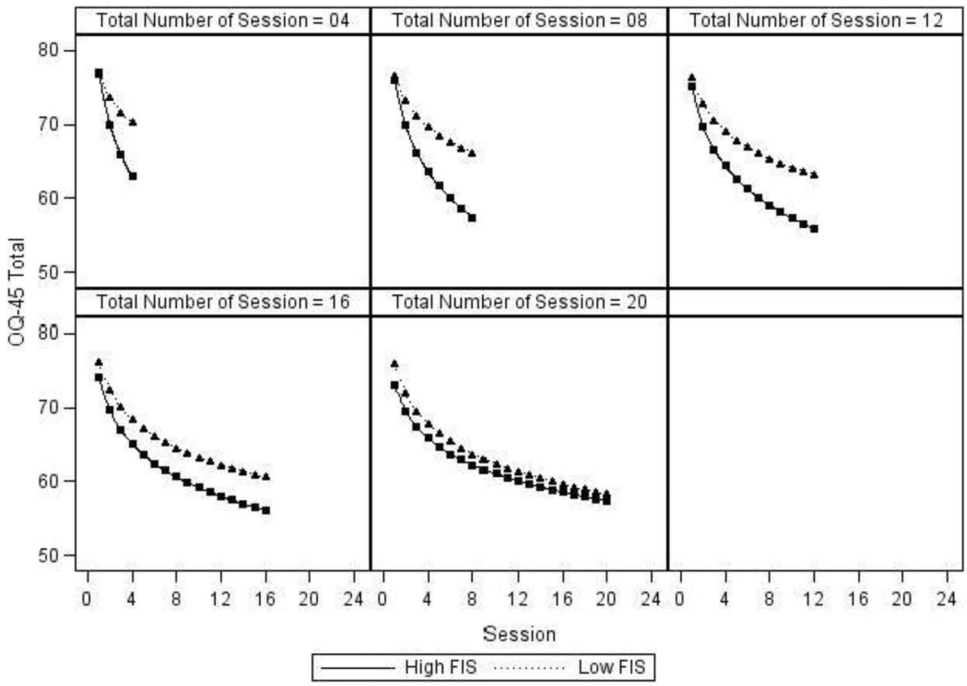

Figure 2. Relationship between therapist FIS and change in client symptoms for different therapy durations. All panels in the figure are model estimates using all data available (and hence are not separate subsamples for each panel). OQ-45 = Outcome Questionnaire-45; FIS = facilitative interpersonal skills.

therapists. This outperformance of high FIS therapists relative to low FIS therapists was attenuated within the 16-session stratum and was absent from the 20-session stratum. Although these findings were limited to the available sample of therapists and client, post hoc power analyses were conducted using the method outlined in Snijders (2005). For the two-level HLM, the power to detect the FIS effect was 56.3%. For the three-level HLM, the power to detect the FIS by log Session effect was 87.0%, the power to detect the number of sessions by log session effect was 60.3%, and the power to detect the FIS by number of sessions by log session effect was 67.4%.

Discussion

This study found that therapist FIS, measured upon entry into graduate training, predicted the outcomes of clients that they treated during their second, third, and fourth years of training. Specifically, clients of higher FIS therapists experienced greater symptom reduction than clients of lower FIS therapists. This is now the third study (see Anderson et al., 2009; Anderson,Crowley, Himawan, Holmberg, & Uhlin, 2015) that has implicated FIS in client outcomes, suggesting that the differential effectiveness of therapists (see Baldwin & Imel, 2013) is at least partly due to therapist interpersonal skills.

The prospective design of this study, where therapist measures were completed 1-year prior to the provision of therapy, provides some evidence that FIS may play a causal role in client outcome. Nonetheless, several aspects of the study were not experimentally controlled. For instance, in this naturalistic setting (i.e., a training clinic), therapy length varied widely. We attempted to account for this by incorporating therapy duration into our hierarchical models (see Baldwin et al., 2009). These analyses indicated that the relation between therapist FIS and client improvement did not remain constant across the various treatment durations. Although higher FIS therapists were more effective than lower FIS therapists for most therapy durations, this effect was (a) particularly pronounced for the shorter therapy durations (e.g., $\leq$ 8 sessions) and (b) virtually nonexistent for the longer-term therapies (e.g., $>$16 sessions). These two findings will be discussed in reverse order.

In light of the latter finding, it might be tempting to conclude that FIS does not affect client progress in long-term therapy. There are two critical points, however, to be made in this context. First, as treatment duration increased, the client sample size also decreased. For instance, approximately 52.3% of clients terminated by the eighth session (and therefore did not contribute outcome data at future sessions), and about 75.7% of clients terminated by Session 15. Therefore, the lack of a FIS-outcome association in long-term therapy only applies for a minority of clients in this study. Second, the duration of therapy is likely an indicator of different client populations because the harder-to-treat clients tend to stay in therapy for a longer period of time (Baldwin et al., 2009; Barkham et al., 2006; DeRubeis, Gelfand, German, Fournier, & Forand, 2014; Stiles, 2013). In other words, while typical clients achieve a *good-enough* level of change and then discontinue therapy (Barkham et al., 2006), the intractable clients experience little to no improvement and thus tend to remain in therapy. Assuming that some clients in our sample were intractable, it stands to reason that these clients would not have responded to high FIS, or perhaps to any intervention for that matter (DeRubeis

et al., 2014), and would consequently be overrepresented in longer-term therapies. In this way, the lack of an FIS-outcome association in long-term therapy may simply be an artifact of intractable clients.

The particularly robust effect of FIS over shorter treatment durations (e.g., ≤8 sessions) was a notable finding from this study. Although the existence of sudden gains is difficult to deny, there are differing and controversial interpretations about what accounts for these sudden gains. For instance, some have suggested that specific technical factors might play a role in sudden gains (e.g., see Aderka, Nickerson, Bøe, & Hofmann, 2012). Given the present findings, it's reasonable to consider to what extent, as well as how, therapist FIS and similar common relationship variables might account for these rapid improvements. Clients' rapid response to high FIS therapists is consistent with a large body of literature indicating that sudden gains in therapy are at least partially attributable to therapist factors and other nonspecific effects (Bohn et al., 2013; Busch, Kanter, Landes, Kohlenberg, 2006; Ilardi & Craighead, 1994; Kelly, Cyranowski, & Frank, 2007). One reason for implicating nonspecific influences, if not common relational influences, is because specific treatment techniques are often not introduced until several sessions into many treatments (e.g., Ilardi & Craighead, 1994). If a treatment rationale, in isolation of any associated techniques, can generate hope and positive expectations, then a therapist who is particularly persuasive and compelling might magnify this early treatment response more than a therapist who is less persuasive (Wampold & Imel, 2015). Forand and Derubeis (2013) noted that clients entering therapy are often eager for relief, and so they may be receptive in the therapist's initial expressions of hope, warmth and understanding. This facilitative environment, according to Forand and DeRubeis (2013), is likely "soothing" (p. 7) for a distressed person and may precipitate a rapid early response.

The present findings would seem to implicate the role of common factors because FIS is, at least partly, a composite of common relational skills (e.g., warmth, alliance capacity, empathy, persuasion). However, it remains unclear which individual skills/processes had the greatest effect on client improvement. For instance, clients may have responded favorably to the therapist's empathic understanding, a convincing treatment rationale, and/or through hope and expectations. It could also be that therapists were skilled in more selective expression of these skills through responsiveness to the particular momentary needs of the client (Stiles, Honos-Webb, & Surko, 1998). Similarly, high FIS therapists may have been effective, not through specific relational actions, but indirectly via a highly facilitative environment, wherein some clients were more likely than others to experience change events (e.g., via insight, corrective experiences). Clearly, future research is needed to understand the specific processes and mechanisms whereby therapist FIS influences clients.

Although the prospective measurement of therapist FIS (at least 1 year prior to the therapists' provision of psychotherapy) is a strength of this study, it is unclear if, or by what degree, therapists' skills (e.g., FIS) might have changed from the time of measurement to the provision of therapy. In this regard, future research should examine the test–retest reliability of therapist FIS.

The naturalistic nature of our study also came at the expense of experimental control. For instance, clients were not randomly assigned to therapists, which may have introduced confounds. Clients in this study were assigned to therapists based on routine procedures of the training clinic, and so it is possible that perceptions of therapist skills influenced the assignment of clients. For example, it's conceivable that client assignments could have been influenced by clinic staff perceptions of which therapists seemed most interpersonally skilled. Our inability to rule out this explanation is a limitation of this study. It should be noted in this context, that we investigated some client- and therapist-level demographics, and these demographics did not predict client outcomes.

Another limitation is that we did not control the number of clients assigned to each therapist; unfortunately, 27.7% of therapists in the current study saw only one client during the data collection period. Although smaller client to therapist ratios minimize the inflation of Type I error rates, the use of only one client per therapist produces experimental confounding (Crits-Christoph & Mintz, 1991). The current research should be replicated using a larger client to therapist ratio.

In addition, the observed results may be reflective of our therapist (trainee) sample; therapists in this study were in the initial phase of their clinical development. Thus, therapists may have struggled to form strong therapeutic relationships while simultaneously delivering techniques that contributed to purposive and collaborative work (Hatcher & Barends, 2006). It is possible that a different pattern of results would have emerged with a more representative sample of therapists, who had additional experience integrating relational skills and specific treatment components into their work.

This intersection between therapist FIS and technical skills, while not directly examined in our study, is a topic worthy of investigation. In this regard, it is important to note that relational skills and technical skills are intertwined and mutually supportive (Bedi, Davis, & Williams, 2005; Hatcher & Barends, 2006; MacFarlane, Anderson, & McClintock, in press; Norcross, 2011). It could be that the warmth and persuasiveness of high FIS therapists augments the effects of their specific treatment techniques (e.g., cognitive restructuring).

Alternatively, high FIS therapists might naturally over rely on their interpersonal skills at the expense of their technical skills. If so, clients of high FIS therapists may have little opportunity to learn and master specific treatment techniques (see Forand & DeRubeis, 2013), which conceivably could undermine the effectiveness of high FIS therapists. Low FIS therapists, on the other hand, may actually compensate for their poor relational skills by employing effective techniques that foster hope and therapeutic progress. In support of this idea, Castonguay et al. (1996) found that therapists increased their adherence to treatment techniques (e.g., cognitive restructuring) to correct relational problems.

Finally, research is needed to examine client characteristics and process variables that might moderate response to therapist FIS; high FIS is presumably more effective for some clients, and in some contexts, than it is for others. Lambert, Garfield, and Bergin (2004) asserted, "When symptoms are not too severe, they seem to respond to the influence of common factors that facilitate change. Greater severity, however, tests the limits of the common factors." (p. 810). In support of Lambert,

Garfield, and Bergin's (2004) assertion, recent evidence indicates that relational effects are moderated by the client's history and severity of symptoms (Lorenzo-Luaces, DeRubeis, & Webb, 2014). Thus, although the FIS of therapists might generally promote healing and well-being, the effects of these skills may ultimately be dependent on a rich web of client, relational, and technical factors that shape that particular therapy experience. More broadly, understanding these variables has implications for selection of students in psychotherapy training, identification of clinical competencies, as well as the development of psychotherapy expertise. As Tracey, Wampold, Lichtenberg, and Goodyear (2014) noted, FIS is one variable that has been identified that characterizes effective therapists and potentially could have significant implications for psychotherapy training and expertise. What's needed even more are studies that will link differences in therapist characteristics and therapist effects in client outcomes to empirically based therapist actions. Knowing how effective therapists express their skills could ultimately anchor judgments of therapist competencies and expertise to therapists' client outcomes.

References

Aderka, I. M., Nickerson, A., Bøe, H. J., & Hofmann, S. G. (2012). Sudden gains during psychological treatments of anxiety and depression: A meta-analysis. *Journal of Consulting and Clinical Psychology, 80,* 93–101. http://dx.doi.org/10.1037/a0026455

Anderson, T., Crowley, M. E., Himawan, L., Holmberg, J. K., & Uhlin, B. D. (2015). Therapist facilitative interpersonal skills and training status: A randomized clinical trial on alliance and outcome. *Psychotherapy Research,* Advance online publication. http://dx.doi.org/10.1080/10503307.2015.1049671

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. http://dx.doi.org/10.1002/jclp.20583

Baldwin, S. A., Berkeljon, A., Atkins, D. C., Olsen, J. A., & Nielsen, S. L. (2009). Rates of change in naturalistic psychotherapy: Contrasting dose-effect and good-enough level models of change. *Journal of Consulting and Clinical Psychology, 77,* 203–211. http://dx.doi.org/10.1037/a0015235

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* (pp. 258–297). New York, NY: Wiley.

Barkham, M., Connell, J., Stiles, W. B., Miles, J. N., Margison, F., Evans, C., & Mellor-Clark, J. (2006). Dose-effect relations and responsive regulation of treatment duration: The good enough level. *Journal of Consulting and Clinical Psychology, 74,* 160–167.

Bedi, R. P., Davis, M. D., & Williams, M. (2005). Critical incidents in the formation of the therapeutic alliance from the client's perspective. *Psychotherapy: Theory, Research, & Practice, 42,* 311–323. http://dx.doi.org/10.1037/0033-3204.42.3.311

Beutler, L. E., Machado, P. P. P., & Neufeldt, S. (1994). Therapist variables. In S. L. Garfield & A. E. Bergin (Eds.), *Handbook of psychotherapy and behavior change* (4th ed., pp. 259–269). New York, NY: Wiley.

Blatt, S. J., Zuroff, D. C., Quinlan, D. M., & Pilkonis, P. (1996). Interpersonal factors in brief treatment of depression: Further analyses of the NIMH Treatment of Depression Collaborative Research Program. *Journal of Consulting and Clinical Psychology, 64,* 162–171. http://dx.doi.org/10.1037/0022-006X.64.1.162

Bohart, A. C., & Tallman, K. (2010). Clients: The neglected common factor in psychotherapy. In B. L. Duncan, S. D. Miller, B. E. Wampold, & M. A. Hubble (Eds.), *The heart and soul of change: Delivering what works in therapy* (2nd ed., pp. 83–111). Washington, DC: American Psychological Association. http://dx.doi.org/10.1037/12075-003

Bohn, C., Aderka, I. M., Schreiber, F., Stangier, U., & Hofmann, S. G. (2013). Sudden gains in cognitive therapy and interpersonal therapy for social anxiety disorder. *Journal of Consulting and Clinical Psychology, 81,* 177–182. http://dx.doi.org/10.1037/a0031198

Boswell, J. F., Gallagher, M. W., Sauer-Zavala, S. E., Bullis, J., Gorman, J. M., Shear, M. K., . . . Barlow, D. H. (2013). Patient characteristics and variability in adherence and competence in cognitive-behavioral therapy for panic disorder. *Journal of Consulting and Clinical Psychology, 81,* 443–454. http://dx.doi.org/10.1037/a0031437

Bryk, A., & Raudenbush, S. W. (1992). Hierarchical linear models for social and behavioral research: *Applications and data analysis methods*. Newbury Park, CA: Sage.

Busch, A. M., Kanter, J. W., Landes, S. J., & Kohlenberg, R. J. (2006). Sudden gains and outcome: A broader temporal analysis of cognitive therapy for depression. *Behavior Therapy, 37,* 61–68. http://dx.doi.org/10.1016/j.beth.2005.04.002

Castonguay, L. G. (2013). Psychotherapy outcome: An issue worth revisiting 50 years later. *Psychotherapy: Theory, Research, & Practice, 50,* 52–67. http://dx.doi.org/10.1037/a0030898

Castonguay, L. G., Goldfried, M. R., Wiser, S., Raue, P. J., & Hayes, A. M. (1996). Predicting the effect of cognitive therapy for depression: A study of unique and common factors. *Journal of Consulting and Clinical Psychology, 64,* 497–504. http://dx.doi.org/10.1037/0022-006X.64.3.497

Crits-Christoph, P., & Mintz, J. (1991). Implications of therapist effects for the design and analysis of comparative studies of psychotherapies. *Journal of Consulting and Clinical Psychology, 59,* 20–26. http://dx.doi.org/10.1037/0022-006X.59.1.20

Crits-Christoph, P., Tu, X., & Gallop, R. (2003). Therapists as fixed versus random effects-some statistical and conceptual issues: A comment on Siemer and Joormann (2003). *Psychological Methods, 8,* 518–523. http://dx.doi.org/10.1037/1082-989X.8.4.518

DeRubeis, R. J., Brotman, M. A., & Gibbons, C. J. (2005). A conceptual and methodological analysis of the nonspecific argument. *Clinical Psychology: Science and Practice, 12,* 174–183. http://dx.doi.org/10.1093/clipsy.bpi022

DeRubeis, R. J., Gelfand, L. A., German, R. E., Fournier, J. C., & Forand, N. R. (2014). Understanding processes of change: How some patients reveal more than others-and some groups of therapists less-about what matters in psychotherapy. *Psychotherapy Research, 24,* 419–428. http://dx.doi.org/10.1080/10503307.2013.838654

Dinger, U., Strack, M., Leichsenring, F., Wilmers, F., & Schauenbirg, H. (2008). Therapist effects on outcome and alliance in inpatient psychotherapy. *Journal of Clinical Psychology, 64,* 344–354.

Elliott, R., Bohart, A. C., Watson, J. C., & Greenberg, L. S. (2011). Empathy. *Psychotherapy, 48,* 43–49. http://dx.doi.org/10.1037/a0022187

Farber, B. A., & Doolin, E. M. (2011). Positive regard and affirmation. In J. Norcross (Ed.), *Psychotherapy relationships that work: Evidence-Based Responsiveness* (2nd ed., pp. 168–186). Oxford, UK: Oxford University Press. http://dx.doi.org/10.1093/acprof:oso/9780199737208.003.0008

Forand, N. R., & Derubeis, R. J. (2013). Pretreatment anxiety predicts patterns of change in cognitive behavioral therapy and medications for depression. *Journal of Consulting and Clinical Psychology, 81,* 774–782. http://dx.doi.org/10.1037/a0032985

Hatcher, R. L., & Barends, A. W. (2006). How a return to theory could help alliance research. *Psychotherapy, 43,* 292–299. http://dx.doi.org/10.1037/0033-3204.43.3.292

Horvath, A. O., Del Re, A. C., Flückiger, C., & Symonds, D. (2011). Alliance in individual psychotherapy. *Psychotherapy: Theory, Research, & Practice, 48,* 9–16. http://dx.doi.org/10.1037/a0022186

Huppert, J. D., Weiss, K. A., Lim, R., Pratt, S., & Smith, T. E. (2001). Quality of life in schizophrenia: Contributions of anxiety and depression. *Schizophrenia Research, 51,* 171–180. http://dx.doi.org/10.1016/S0920-9964(99)00151-6

Ilardi, S. S., & Craighead, W. E. (1994). The role of nonspecific factors in cognitive-behavior therapy for depression. *Clinical Psychology: Science and Practice, 1,* 138–155. http://dx.doi.org/10.1111/j.1468-2850.1994.tb00016.x

Imel, Z. E., Baer, J. S., Martino, S., Ball, S. A., & Carroll, K. M. (2011). Mutual influence in therapist competence and adherence to motivational enhancement therapy. *Drug and Alcohol Dependence, 115,* 229–236. http://dx.doi.org/10.1016/j.drugalcdep.2010.11.010

Kelly, M. A. R., Cyranowski, J. M., & Frank, E. (2007). Sudden gains in interpersonal psychotherapy for depression. *Behaviour Research and Therapy, 45,* 2563–2572. http://dx.doi.org/10.1016/j.brat.2007.07.007

Kim, D. M., Wampold, B. E., & Bolt, D. M. (2006). Therapist effects in psychotherapy: A random-effects modeling of the National Institute of Mental Health Treatment of Depression Collaborative Research Program data. *Psychotherapy Research, 16,* 161–172. http://dx.doi.org/10.1080/10503300500264911

King, M. F., & Bruner, G. C. (2000). Social desirability bias: A neglected aspect of validity testing. *Psychology & Marketing, 17,* 79–103. http://dx.doi.org/10.1002/(SICI)1520-6793(200002)17:2<79::AID-MAR2>3.0.CO;2-0

Lambert, M. J., Garfield, S. L., & Bergin, A. E. (2004). Overview, trends, and future issues. In M. J. Lambert (Ed.), *Bergin and Garfield's handbook of psychotherapy and behavior change* (5th ed., pp. 805–821). New York, NY: Wiley.

Lambert, M. J., Lunnen, K., Umphress, V., Hansen, N., & Burlingame, G. M. (1994). *Administration and scoring manual for the Outcome Questionnaire (OQ-45.1).* Salt Lake City, UT: IHC Center for Behavioral Healthcare Efficacy.

Loewenstein, J. S., & Lerner, J. S. (2003). The role of affect in decision making. In R. J. Davidson, K. R. Scherer, & H. H. Goldsmith (Eds.), *Handbook of affective science* (pp. 619–642). New York, NY: Oxford University Press.

Lorenzo-Luaces, L., DeRubeis, R. J., & Webb, C. A. (2014). Client characteristics as moderators of the relation between the therapeutic alliance and outcome in cognitive therapy for depression. *Journal of Consulting and Clinical Psychology, 82,* 368–373. http://dx.doi.org/10.1037/a0035994

MacFarlane, P., Anderson, T., & McClintock, A. S. (in press). The early formation of the working alliance from the client's perspective: A qualitative study. *Psychotherapy: Theory, Research, & Practice*.

Norcross, J. C. (2011). *Psychotherapy relationships that work: Evidence-Based Responsiveness* (2nd ed.). Oxford, UK: Oxford University Press. http://dx.doi.org/10.1093/acprof:oso/9780199737208.001.0001

Ogles, B. M. (1996). Assessing outcome in practice. *Journal of Mental Health, 5,* 35–46. http://dx.doi.org/10.1080/09638239650037171

Okiishi, J. C., Lambert, M. J., Nielsen, S. L., & Ogles, B. M. (2003). Waiting for supershrink: An empirical analysis of therapist effects. *Clinical Psychology & Psychotherapy, 10,* 361–373. http://dx.doi.org/10.1002/cpp.383

Riggio, R. E. (1986). Assessment of basic social skills. *Journal of Personality and Social Psychology, 51,* 649–660. http://dx.doi.org/10.1037/0022-3514.51.3.649

Riggio, R. E. (1989). *Manual for the social skills inventory*. Palo Alto, CA: Consulting Psychologists Press.

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

Singer, J. (1998). Using SAS PROC MIXED to fit multilevel models, hierarchical models, and individual growth models. *Journal of Educational and Behavioral Statistics, 24,* 323–355. http://dx.doi.org/10.2307/1165280

Snijders, T. A. B. (2005). Power and sample size in multilevel linear models. In B. S. Everitt & D. C. Howell (Eds.), *Encyclopedia of statistics in behavioral science* (vol. 3, pp. 1570–1573). Chicester, UK: Wiley. http://dx.doi.org/10.1002/0470013192.bsa492

Stiles, W. B. (2013). The variables problem and progress in psychotherapy research. *Psychotherapy, 50,* 33–41. http://dx.doi.org/10.1037/a0030569

Stiles, W. B., Honos-Webb, L., & Surko, M. (1998). Responsiveness in psychotherapy. *Clinical Psychology: Science and Practice, 5,* 439–458. http://dx.doi.org/10.1111/j.1468-2850.1998.tb00166.x

Stone, A. A., Turkkan, J. S., Bachrach, C. A., Jobe, J. B., Kurtzman, H. S., & Cain, V. S. (2000). *The science of Self-report*. Mahwah, NJ: Lawrence Erlbaum Associates.

Tourangeau, R. (2009). Remembering what happened: Memory errors and survey reports. In A. A. Stone, J. S. Turkkan, C. A. Bachrach, J. B. Jobe, H. S. Kurtzman, & V. S. Cain (Eds.), *The science of self-report* (pp. 29–48). Mahwah, NJ: Lawrence Erlbaum Associates.

Tracey, T. J. G., Wampold, B. E., Lichtenberg, J. W., & Goodyear, R. K. (2014). Expertise in psychotherapy: An elusive goal? *American Psychologist, 69,* 218–229. http://dx.doi.org/10.1037/a0035099

Wampold, B. E., & Bolt, D. M. (2006). Therapist effects: Clever ways to make them (and everything else) disappear. *Psychotherapy Research, 16,* 184–187. http://dx.doi.org/10.1080/10503300500265181

Wampold, B. E., & Brown, G. S. (2005). Estimating therapist variability: A naturalistic study of outcomes in managed care. *Journal of Consulting and Clinical Psychology, 73,* 914–923.

Wampold, B. E., & Imel, Z. E. (2015). *The great psychotherapy debate: The evidence for what makes psychotherapy work*. New York, NY: Routledge.

Received January 5, 2015
Revision received July 31, 2015
Accepted September 4, 2015 ■