Facilitative interpersonal skills in benign *versus* challenging therapy situations in trainee therapists: a pilot study

Kim de Jong,[1] Johanna Wilkens,[1] Timothy Anderson,[2] Kane Steggles[3]

[1]Clinical Psychology Unit, Institute of Psychology, Leiden University, the Netherlands; [2]Department of Psychology, Ohio University, Athens, OH, USA; [3]School of Psychology, University of Southampton, UK

ABSTRACT

Therapists' responses to challenging therapy situations on the Facilitative Interpersonal Skills (FIS) performance task are a significant predictor of therapists' differences in treatment outcomes. The aim of this study was to assess whether the complexity of the therapy situation influenced the facilitative interpersonal skills of trainees. Trainee therapists (*n*=46) participated in an experiment in which they responded to a set of challenging and benign (*i.e.*, non-challenging) video vignettes of therapy situations of the FIS performance task. Their responses were video recorded and coded by four independent raters. Results showed that trainees scored significantly higher on the FIS performance task responding to benign therapy situations than responding to challenging situations. This is the first study to investigate difficulty of therapy situations as a potential predictor or trainees interpersonal skills. Further research is needed to replicate these results in a larger sample.

Key words: facilitative interpersonal skills, therapist effects, therapist training.

Introduction

Psychotherapy research consistently suggests that some therapists are better in eliciting adaptive change and enhancing treatment outcomes with their clients than others (Okiishi *et al.*, 2003; Wampold & Bolt, 2006). These differences in the effectiveness of therapists, referred to as therapist effects, account for approximately 7% of the variance in patient outcomes (Schiefele *et al.*, 2017). Although the number of studies on therapist effects has increased in the past decade, only few consistent predictors have been found. Two systematic reviews of the therapist effects literature concluded that the evidence for what characterizes effective therapists is nascent (Heinonen & Nissen-Lie, 2020; Wampold & Owen, 2021). The exception is research on therapists' observed facilitative interpersonal skills, which consistently predicts differences between therapists in outcomes (*i.e.*, symptom reduction), even up to five years later (Anderson, Mc-Clintock, *et al.*, 2016; Anderson *et al.*, 2009; Anderson *et al.*, 2007; Schöttke *et al.*, 2016). Therapists' facilitative interpersonal skills are generally measured using the Facilitative Interpersonal Skills (FIS) performance task (Anderson *et al.*, 2007), which uses video vignettes of challenging therapy situations, typically alliance ruptures (*e.g.*, angry patient, passive aggressive patient).

Facilitative interpersonal skills (FIS) can be defined as skills that can help to facilitate positive change in individuals in emo-

tional and psychological distress (Anderson *et al*., 2019). The FIS performance task aims to assess the following core interpersonal skills: verbal fluency, emotional expression, persuasiveness, warmth/positive regard, hopefulness, empathy, alliance bond capacity, and alliance-rupture-repair responsiveness. The task entails responding to a series of interpersonally challenging therapy situations with different clients, the responses are video recorded and then assessed by trained raters. Previous studies found that therapists with high FIS scores showed better treatment outcomes and better therapeutic alliance compared to those with low FIS scores based on clients' self-reported measures and independent clinical assessments (Anderson, Crowley, *et al*., 2016; Anderson, McClintock, *et al*., 2016; Anderson *et al*., 2009).

Although the FIS task has been a consistent predictor of therapists' treatment outcomes, it is still somewhat unclear what the underlying mechanism of action is. In their recent review, Allen and colleagues (2023) state that when a participant must immediately respond to a challenging therapy situation, they are more likely to experience the heightened emotion and urgency that can accompany challenging therapy interactions. However, video simulations also cause limitations, since the therapist does not have a relationship with the patient in the video and is acting in a test situation.

We hypothesize that the challenging nature of the FIS task is one of the reasons it has been so consistent in predicting treatment outcomes. Previous research has demonstrated that therapist effects are more pronounced in patients with higher symptom severity at intake, who are often more challenging in treatment, than in patients with milder symptoms (Saxon & Barkham, 2012). It seems that treating more challenging patients may require therapists to utilize their skills more effectively, which is in line with the idea that therapy quality matters most in more challenging patients (DeRubeis *et al*., 2014).

The aim of this pilot study is to explore whether the difficulty of the cases influences therapists' scores on the FIS task. For this purpose, we developed a new set of therapy situation vignettes for the FIS task, in which benign (*i.e.*, non-challenging) therapy situations were displayed. Therapists' FIS scores could go in both directions. They could either have lower FIS scores in the challenging condition, because the therapy situations are more difficult to respond to. Or, they could have higher FIS scores, because they would need to utilize their interpersonal skills more in the challenging therapy situation. For example, in a benign situation there may not be a need to use language that strengthens the alliance, which may lead to a lower score on the alliance bond capacity domain.

To our knowledge, this is the first study that aims to investigate the working mechanisms of the FIS performance task. The research question was whether there was a difference in FIS scores in responding to the benign and the challenging therapy situations in a sample of trainee therapists. Since this is the first study using difficulty as a variable in assessing FIS, we do not have a specific hypothesis regarding the direction of the relationship.

Methods

Design

The study used a within-subject design, using a fixed order of conditions (benign and challenging). Given this study was a pilot study, there was no counterbalancing of conditions to restrict variability.

Participants

A total of 46 trainee therapists took part in the study. Participants were enrolled in the one-year English clinical psychology master program at Leiden University and were invited to take part in the study within the first month of starting their clinical master program via e-mail. They were asked to sign up for the study through the university's online research participation system (SONA). Inclusion criteria were i) enrolment in the Clinical Psychology Master track at Leiden University, and ii) English fluency. Exclusion criteria were related to physiological data being collected as part of the same study (not analysed in this manuscript) and included i) known heart problems, ii) use of beta-blockers, and iii) consumption of caffeinated beverages two hours prior to the experiment, as these all affect physiological measurements of heart rate. Participation was on a voluntary basis. Written informed consent was obtained. Participants received a monetary incentive of ten euros in return for their participation. The experiment was granted ethical approval by the Psychology Research Ethics Committee (CEP) at Leiden University prior to recruitment (ethical approval number: CEP17-1222/434).

Three participants were missing completely at random: one participant had to be excluded due for technical difficulties during the FIS task and two participants had missing grades, because they did not study in our department. Thus, the final sample consisted of 43 participants. The sample consisted to 95.3% of females and had an age range from 21 to 47 with a mean of 25.4 (SD=5.02). Regarding nationality, 34.9% were German, 20.9% were Dutch, 14.0% were Greek, 7.0% were from the United Kingdom and 23% had another nationality.

Facilitative Interpersonal Skills performance task (manipulation)

The FIS performance task was designed by Anderson and colleagues (2007) based on cases from the Vanderbilt II psychotherapy study (Strupp, 1993) where out of 80 cases they chose the most challenging seven cases to represent a variety of interpersonally challenging situations of clients on both the friendly-hostile and submissive-dominant domains on the interpersonal circumplex. The FIS task had shown good internal consistency, validity, and an inter-rater reliability that ranges between .70 and .80 (Anderson & Patterson, 2013). In the original task, participants were exposed to seven video vignettes (1-2 minutes each) of therapy situations. Each vignette starts with an introduction to the client by a narrator, followed by a section where the client talks to the therapist as if it were a real therapy situation. The vignette ends with a response from the client that is interpersonally challenging for the therapist. Then there is text on screen to inform the participant that it is their turn to talk

The original FIS task only included interpersonally challenging situations directed at the therapist (*e.g.*, client is angry at the therapist). In order to test whether the challenging nature of these situations was vital for the predictive value of the FIS performance task, seven additional vignettes were developed (De Jong *et al*., 2018). These so-called *benign* vignettes were based on the same therapy session that the original vignettes were developed from but used a section of the session in which the client describes an interpersonal situation (*e.g.*, going home for Christmas and meeting old friends) to the therapist. The main difference between the original challenging vignettes and the new benign vignettes is that in the benign vignettes the interper-

sonal situation at hand were less directed towards the therapist and presumably less challenging for the therapist. The benign vignettes were filmed working with actors for the purpose of the project. Similar to the challenging condition, each vignette starts with introducing the client by a narrator followed by a section where the client talks about an interpersonal problem.

After each vignette, participants were asked to respond to the client as if they were the client's therapist. The responses were video recorded and scored by four independent raters using the FIS rating manual (Anderson & Patterson, 2013). All raters were trained by the first and last author, who consulted with the third author when necessary. Training involved instructions on coding the eight domains of FIS, discussing raters' own potential biases, practicing on two sets of practice videos provided by the third author and calibration meetings to get the coders to discuss questions about the coding.

Instruments

FIS coding manual

The Facilitative Interpersonal Skills of the participants were rated based on the FIS coding manual (Anderson & Patterson, 2019). The answers were rated on a five-point scale, the scoring of which is based on descriptions of observable behaviour. For example, a score of 1 on the domain of Verbal Fluency would be assigned if *the participant has great difficulty verbalizing a response, reflecting a clear avoidance or anxiety* and a score of 5 is assigned when *the participant is at great ease and communicates ideas with no anxiety, reflecting a desire to approach the other*. The FIS task has the following domains: i) verbal fluency (VF), ii) hope & positive expectations (HPE), iii) persuasiveness (PE), iv) emotional expression (EE), v) warmth, acceptance & understanding (WAU), vi) empathy (EMP), vii) alliance bond capacity (ABC), and viii) alliance rupture repair responsiveness (ARRR)/problem responsiveness (PR). The last item was rated as ARRR if the situation was challenging, and PR if the situation was benign. The PR domain was developed to replace the Alliance Rupture Repair Responsiveness in the benign clips, because alliance ruptures did not exist in the benign situations. Both measures aim at assessing the therapist's ability to respond to the client's specific problem, whether it relates to the alliance or not, in a facilitative way. The FIS score is computed by taking the mean of the eight domains.

Demographic variables

The demographic data such as the participant's sex, age, country of birth, personal and/or (semi)professional experience with mental disorders and English proficiency were collected for all participants.

Grades

Clinical grades were the mean grade of three clinical courses as part of the master program (Basic Therapeutic Skills, Cognitive Behavioural Interventions, and Clinical Interviewing and Assessment) and the clinical internship (equivalent of 4 months full-time) or Internal Practice Internship (IPI; equivalent of 2 months full-time). Each of the courses consisted of weekly workgroup meetings in which students practiced therapy skills in a simulated therapy setting using classmates or actors. Clinical internships were in mental health care organizations in a variety of settings, whereas the IPI consisted of counselling (5

sessions) of clients with mild to moderate symptom severity. Academic grades were the mean of two theoretical courses (research methods, clinical interventions theory exam) and the master thesis. Missing values were replaced by the mean of the available grades if not more than one grade was missing.

Procedure

The study was part of a bigger project that aimed at investigating the underlying working mechanisms of FIS. All data were collected by a team of clinical psychology/clinical and health research master students and a research intern. Participants were invited individually to a room where they were seated in front of a computer to perform the FIS task as they were being recorded. All participants watched and responded to the benign cases first and then to the challenging cases. After finishing the task, participants were debriefed about the aim of the study and were asked to sign a second consent form, that gave permission to access their grades.

Statistical analyses

First, it was assessed whether there was a relationship between FIS and clinical grades, the proxy for outcomes in our study. The manipulation check for distinguishing between the benign and challenging therapy situations was conducted with two-tailed t-tests, and the second research question was answered using a one-sample t-test with the difference in FIS score (benign minus challenging) between the challenging and benign clips as the dependent variable.

Results

Preliminary analyses

Firstly, we examined descriptive statistics, which can be seen in Table 1. In data exploration, one academic grade was identified as a mild outlier, but fitted the natural variation in scores seen in grades and was thus retained. No violations of the assumptions for the regression analyses and one sample t-test were found. Both the FIS challenging and benign scores, academic and the clinical grades were normally distributed. In addition, there were no problems of homoscedasticity, multicollinearity or violating independence in the error terms.

Does trainees performance on the FIS task predict outcomes?

Hierarchical multiple regression was used to examine the extent to which performance on the FIS task predicts clinical grades, while accounting for age, experience with patients, lived experience as a patient, English proficiency, academic grade, and sex as covariates (Table 2).

In the first step of the analysis, FIS was entered, explaining 10.6% of the variance in clinical grades, $F(1, 41)=4.84$, $p=.03$. Following the inclusion of the covariates in Step 2 (age, experience with patients, lived experiences as a patient, English proficiency, academic grade, and sex), the total variance explained by the model increased to 64.6%, $F(7, 35) = 8.89$, $p<.001$. This indicated that the covariates accounted for an additional 57.5% of the variance in clinical grades, $R^2_{change}=.54$, $F_{change}(6, 35)=8.89$, $p<.001$, beyond the influence of FIS. Within the final model encompassing both FIS and covariates, only age and academic

grade were found to be statistically significant predictors of clinical grades (Table 2). This suggests that only age and academic grade exerted significant independent effects on clinical grade outcomes after controlling for the other variables in the model.

Is there a difference in responding to benign *vs* challenging therapy situations?

To address the main research question of whether there are differences in FIS scores between responses to challenging versus benign therapy situations, we first conducted a manipulation check. We compared participants' ratings of difficulty, distress, and concentration during challenging and benign therapy situations (Table 3). The results revealed no significant differences in perceived difficulty and distress between challenging and be-

nign therapy situations, $t(42)=-0.16$, $p=.87$; $t(42)=1.54$, $p=.12$ respectively. However, participants reported significantly higher levels of concentration during challenging therapy situations compared to benign therapy situations, $t(42)=-3.57$, $p<.001$ (two-tailed). The Cohen's d statistic indicated a medium effect size ($d=0.55$). We intended the challenging clips to be more difficult than the benign clips, but it does not seem like that is how participants perceived them. We also see that participants reported being more concentrated during their responses to the challenging clips.

Next, we conducted a one-sample t-test to examine the difference in FIS scores between challenging and benign therapy situations (Table 3). The analysis revealed that FIS scores were significantly lower in response to challenging therapy situations compared to benign therapy situation, $t(42)=2.93$, $p=.005$ (two-

Table 1. Descriptive statistics for study variables (n=43).

Variable	%	Range	Mean (SD)
FIS challenging		2.22-3.95	3.08 (0.38)
FIS benign		2.14-4.19	3.25 (0.46)
Clinical grade		7.38-9.00	8.11 (0.39)
Academic grade		6.17-9.17	7.96 (0.65)
DERS		43.00-131.00	77.58 (20.67)
GSES		23.00-38.00	31.05 (3.32)
English proficiency		23.00-40.00	33.26 (5.03)
Experience with patients	72		
Lived experience as patient	42		

FIS, Facilitative Interpersonal Skills; SD, standard deviation; DERS, Difficulties in Emotion Regulation Scale; GSES, Generalized Self-Efficacy Scale.

Table 2. Hierarchical regression results for clinical grade (n=43).

Variables	B	SE	95% CI for B		β	R²	DR²
			LL	UL			
Step 1						0.11	0.11
Constant	7.07***	.48	6.11	8.04			
FIS	0.34*	.15	0.03	0.65	.33*		
Step 2						0.56	0.54
Constant	4.10***	.56	2.98	5.23			
FIS	0.07	.11	-0.16	0.30	.07		
Age	0.03**	.01	0.01	0.05	.37**		
Sex	0.32	.21	-0.12	0.75	.17		
Experience with patients	0.00	.10	-0.19	0.20	.00		
Lived experiences as a patient	-0.06	.09	-0.24	0.12	-.08		
English proficiency	0.02	.01	-0.00	0.04	.22		
Academic grade	0.28***	.08	0.13	0.43	.45***		

SE, standard error; CI, confidence interval; LL, lower limit; UL, upper limit; FIS, Facilitative Interpersonal Skills. *p<.05. **p<.01. ***p<.001.

Table 3. Difference between benign and challenging cases in difficulty, distress, concentration, and Facilitative Interpersonal Skills score (n=43).

Variable	Benign, M (SD)	Challenging, M (SD)	Mean difference	95% CI for mean difference		Cohen's d
				LL	UL	
Difficulty	53.00 (18.17)	53.34 (23.08)	0.35	-4.65	3.95	0.03
Distress	43.75 (19.11)	40.74 (22.06)	3.01	-0.94	6.97	0.24
Concentration	73.63 (13.79)	77.00 (12.10)	-3.37***	-5.27	-1.47	0.55
FIS scores	3.25 (0.45)	3.08 (0.38)	0.17**	0.05	0.29	0.45

CI, confidence interval, M, mean; SD, standard deviation; LL, lower limit, UL, upper limit; FIS, Facilitative Interpersonal Skills. *p<.05. **p<.01. ***p<.001 (two-sided).

tailed). The Cohen's d statistic indicated a small to medium effect size (d=0.45).

Discussion

The aim of this study was to assess a potential working mechanism for the predictive nature of the Facilitative Interpersonal Skills performance task, which is one of the more robust predictors of therapist differences in treatment outcomes. We hypothesized that the challenging nature of the task may partially explain its predictive utility. Given that therapist effects are more pronounced in more severe, and likely more challenging cases (Saxon & Barkham, 2012), it stands to reason that more complex stimulus clips may further distinguish high and low performing therapists within the FIS task. We tested this hypothesis in a sample of master student trainees. The first question was whether the FIS predicted outcomes in trainees, in which we used clinical grade as a proxy for treatment outcomes. FIS was indeed predictive of clinical grades, but after controlling for covariates, this was no longer the case. In particular, academic grades were a stronger predictor of outcomes, suggesting that general aptitude might be more important in predicting clinical grades than interpersonal skills. One reason for that might be that the clinical grades are not solely based on performance in patient-therapist situations, but also on written reports that reflect on their own learning objectives and performance. Thus, clinical grades may not be the best proxy for treatment outcomes in trainees, which is in line with previous studies (Costanzo & Philpott, 1986; Loo, 1979).

The second research question was whether there was a difference in facilitative interpersonal skills in responding to challenging vs. benign therapy situations. The results indicated that trainees scored significantly lower on the FIS in challenging than in benign therapy situations. Trainees did not perceive the challenging situations to be more difficult than the benign cases on average, but did indicate that they were experiencing higher levels of concentration while responding to them. These results might be an indication that trainees have more trouble utilizing their interpersonal skills in more challenging therapy situations. One reason could be that since they had typically not worked with patients before, and scenarios in the training program were possibly more benign than challenging. As such, they may also simply not have that much experience with more challenging situations. Alternatively, it could also be that the challenging cases were outside their zone of proximal development (not too difficult, not too easy; Rousmaniere *et al*, 2017), causing overwhelm. For example, we observed that some participants seemed overwhelmed by the task and broke protocol by starting to talk about the case, rather than to the patient. Differences between trainees in performance on the FIS task might be determined in part by differences in emotion regulation skills of the trainee – some trainees may be better able to regulate the emotions that are evoked by the challenging therapy situations in themselves than others. Similarly, challenging therapy situations might be evoking stronger responses of countertransference than benign therapy situations. In the case of countertransference or difficulties in emotion regulation, it is likely more difficult for trainees to utilize their facilitative interpersonal skills, even if they do possess them.

It is also possible that the difference in scores is a result of bias in rating. Because it was needed to know the context of the therapy situation to code the responses, it is possible that coders scored participants structurally lower in the challenging situations than in the benign ones. Result would need to be replicated to get a better idea of whether this effect is found in other trainees as well, and what would be the response in experienced therapists.

This study has several limitations that are influencing the results. Being a pilot study, the sample size was relatively small, although the sample size is larger than many other studies with therapists or trainees. Still, the analyses are likely underpowered, making it uncertain whether the results will replicate or not. Participants were also predominantly female, and thus the results may not generalize to male trainees and therapists. In addition, we did not counterbalance the design, making it difficult to determine whether order effects might play a role here. Trainees were all offered the benign cases first, giving them the opportunity to practice with less challenging therapy situations. However, this did not result in them performing better in the challenging situations – or at least not better than in the benign situations. It is advisable for future studies to counterbalance conditions and control for potential order effects in that way. Furthermore, participants were not treating patients yet and thus treatment outcomes were not available to assess therapists' effectiveness. We used clinical grades as a proxy for treatment outcomes, but it is unclear to what extent these are related to clinical outcomes. Future studies should investigate this in a sample of trainees who have treatment outcomes in patients.

If these results are replicated in future studies, they may have implications for clinical training. More challenging therapy situations might reduce trainees ability to respond in a facilitative manner, and given that FIS is predictive of treatment outcomes, further emphasis on training that focuses on responding to challenging therapy situations might be warranted. Allen *et al*. (2023) suggest that the FIS task can be emotionally challenging for therapists to encounter, so practice may habituate trainees to these situations, making it easier to respond to them in real life.

Conclusions

This was the first study assessing the underpinnings of the FIS performance task, looking into difficulty as a potential predictor of performance. The newly developed set of benign therapy situations show that trainees indeed respond differently to them, even though participants did not perceive them as more difficult. As a result, difficulty of therapy situations might be factor of interest for future research in trainees.

References

Anderson, T., & Patterson, C. L. (2013). Facilitative interpersonal skills performance analysis rating method. Unpublished coding manual. Department of Psychology, Ohio University, Athens, OH.

Anderson, T., McClintock, A. S., Himawan, L., Song, X., & Patterson, C. L. (2016). A prospective study of therapist facilitative interpersonal skills as a predictor of treatment outcome. *Journal of Consulting and Clinical Psychology, 84*(1), 57-66. https://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(7), 755-768. https://doi.org/10.1002/jclp.20583

Anderson, T., Patterson, C., McClintock, A. S., McCarrick, S. M., Song, X., & The Psychotherapy and Interpersonal Lab Team. (2019). *Facilitative interpersonal skills task and rating manual*. Ohio University, Athens, OH.

Anderson, T., Patterson, C., & Weis, A. (2007). Facilitative interpersonal skills performance analysis rating method. Unpublished coding manual. Department of Psychology, Ohio University, Athens, OH.

Costanzo, M., & Philpott, J. (1986). Predictors of therapeutic talent in aspiring clinicians: A multivariate analysis. *Psychotherapy, 23*(3), 363-369. https://doi.org/10.1037/h0085624

De Jong, K., et al. (2018). Additional video material for the Facilitative Interpersonal Skills task: Benign vignettes. Leiden, Leiden University.

DeRubeis, R. J., Cohen, Z. D., Forand, N. R., Fournier, J. C., Gelfand, L. A., & Lorenzo-Luaces, L. (2014). The Personalized Advantage Index: Translating research on prediction into individualized treatment recommendations. A demonstration. *PLOS ONE, 9*(1), e83875. https://doi.org/10.1371/journal.pone.0083875

Heinonen, E., & Nissen-Lie, H. A. (2020). The professional and personal characteristics of effective psychotherapists: A systematic review. *Psychotherapy Research, 30*(4), 417-432. https://doi.org/10.1080/10503307.2019.1620366

Loo, R. (1979). Role of primary personality factors in the perception of traffic signs and driver violations and accidents. *Accident Analysis and Prevention, 11*(2), 125-127. https://doi.org/10.1016/0001-4575(79)90020-4

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*(6), 361-373.

Saxon, D., & Barkham, M. (2012). Patterns of therapist variability: Therapist effects and the contribution of patient severity and risk. *Journal of Consulting and Clinical Psychology, 80*(4), 535-546. https://doi.org/10.1037/a0028898

Schiefele, A.-K., Lutz, W., Barkham, M., Rubel, J., Böhnke, J., Delgadillo, J., Kopta, M., Schulte, D., Saxon, D., & Nielsen, S. L. (2017). Reliability of therapist effects in practice-based psychotherapy research: A guide for the planning of future studies. *Administration and Policy in Mental Health and Mental Health Services Research, 44*(5), 598-613.

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, 26*(5), 542-551. https://doi.org/10.1080/10503307.2015.1125546

Seedall, R. B. (2019). Review of the book *The Cycle of Excellence: Using Deliberate Practice to Improve Supervision and Training*, by T. Rousmaniere, R. K. Goodyear, S. D. Miller, & B. E. Wampold (Eds.). *Journal of Marital and Family Therapy, 45*(1), 187-188. https://doi.org/10.1111/jmft.12361

Strupp, H. H. (1993). The Vanderbilt Psychotherapy Studies: Synopsis. *Journal of Consulting and Clinical Psychology, 61*(3), 431-433. https://doi.org/10.1037/0022-006X.61.3.431

Wampold, B. E., & Bolt, D. M. (2006). Therapist effects: Clever ways to make them (and everything else) disappear. *Psychotherapy Research, 16*(2), 184-187.

Wampold, B. E., & Owen, J. (2021). Therapist effects: History, methods, magnitude. In M. J. Lambert (Ed.), *Bergin and Garfield's Handbook of Psychotherapy and Behavior Change* (7th ed., pp. 297-326). Wiley.