







Development and validation of the facilitative interpersonal skills scale for clients

João M. Santos¹  | Matilde Barata²  | Sara Rathenau²  |
Inês Amaro²  | Alexandre Vaz³  | Daniel Sousa⁴  |
Margarida Severino⁵ | Mafalda Taveira⁵

¹Centro de Investigação e Estudos de Sociologia (CIES-Iscte), Iscte—Instituto Universitário de Lisboa, Lisbon, Portugal

²ISPA University Clinic, ISPA—University Institute, Lisbon, Portugal

³Department of Clinical Psychology, Sentio Counseling Center, Los Angeles, United States

⁴Clinical and Health Psychology Department, ISPA—University Institute, Lisbon, Portugal

⁵Private Practice, Lisbon, Portugal

Correspondence

João M. Santos, Centro de Investigação e Estudos de Sociologia (CIES-Iscte), Iscte—Instituto Universitário de Lisboa, Lisbon, Portugal.
Email: joao_marques_santos@iscte-iul.pt

Abstract

Objective: Psychotherapy studies have revealed that therapist characteristics are responsible for 5% to 9% of outcome variance. The therapist-facilitative interpersonal skills (FIS) have been shown to predict both alliance and outcomes, indicating that higher FIS therapists are more effective than lower FIS therapists. The current study focused on the development and validation of the FIS-client version (FIS-C) instrument, aimed at collecting the clients' perspectives on relevant therapist characteristics.

Method: The clinical outcomes in routine evaluation—outcome measures, the session rating scale, and the FIS questionnaire—client version were filled out by psychotherapy clients. Exploratory, confirmatory factor, and test-retest analysis were conducted.

Results: Results indicate robust psychometric characteristics, in terms of validity (factorial, convergent, discriminant, and nomological), reliability, and sensitivity.

Conclusion: The validation of the FIS-C represents an important contribution to clinical research and practice, namely to the field of client feedback and therapist expertise.

This is an open access article under the terms of the Creative Commons Attribution-NonCommercial License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited and is not used for commercial purposes.
© 2022 The Authors. *Journal of Clinical Psychology* published by Wiley Periodicals LLC.

KEYWORDS

psychotherapist development, psychotherapy expertise, statistical methodology, test development, therapist effects

1 | INTRODUCTION

The role of the therapist in psychotherapy research has a controversial history. The widespread focus on investigating specific treatment protocols for specific psychological disorders made the study of individual differences among therapists a long-neglected research topic (e.g., Beutler et al., 2004; Castonguay & Hill, 2017). However, despite the historically greater emphasis on randomized controlled trials to understand differences between specific treatment models, research has consistently shown that *therapists* contribute more to therapy outcomes than the type of treatment or level of adherence to a particular protocol (Wampold & Imel, 2015). It seems that therapist effects can explain 5%–9% of the outcome variance (Baldwin & Imel, 2013). These results overshadow the 0%–1% variability attributable to specific treatment models (Miller et al., 2013; Wampold & Imel, 2015).

The existing literature suggests that the therapeutic alliance is one of the best predictors of clinical outcomes identified to date, along with the therapist's ability to convey empathy (Norcross & Lambert, 2019; Norcross & Wampold, 2019; Wampold & Imel, 2015). The therapeutic alliance is described by Bordin (1979) as a collaborative relationship intended to overcome the client's suffering. This relationship encompasses three important aspects: (a) agreement on the goals of the treatment, (b) agreement on the tasks, and (c) the development of a mutual bond between therapist and client. The therapeutic alliance is considered a common factor or a nonspecific factor in psychotherapy. These factors include therapeutic qualities (such as empathy and the therapist's ability to be persuasive and to create compelling expectations) found in therapists and professionals across therapeutic approaches and helping professions (Frank & Frank, 1993). Common factors are subject to various definitions which poses some challenges to controlled experimental research (Anderson & Patterson, 2013). Following the need to measure and operationalize these factors, Anderson and Patterson (2013) created the facilitative interpersonal skills (FIS) rating scale, which assesses several relational therapist skills relevant to therapeutic effectiveness (Anderson & Patterson, 2013). These skills play an important role in the ability to develop and maintain a solid therapeutic alliance and might go beyond the therapist's use of a particular treatment model or technique (Wampold & Imel, 2015).

The FIS rating scale (Anderson & Patterson, 2013) is a performance-based measure of eight therapist interpersonal skills, namely verbal fluency, emotional expression, persuasiveness, warmth/positive regard, hopefulness, empathy, alliance bond capacity, and alliance–rupture–repair responsiveness. This measure was designed for the assessment of the therapist's performance by an external observer. More specifically, it evaluates participants' responses to a performance-based task where they provide therapeutic responses to standard video stimuli that portray challenging psychotherapeutic events. In these clips, one actor/actress plays the role of a client explaining a certain problem to which the participants have to respond in a therapeutic manner as if they were in session. The FIS-in session (FIS-IS) rating scale (Uhlen, 2011) is an adaptation of the original FIS rating scale applied to actual recorded or observed therapy sessions. Similarly to the original FIS, an external coder rates the therapist's performance. Our goal is to create a new FIS scale (the FIS-client version [FIS-C]) where clients evaluate their therapists' performance.

Anderson et al. (2009) found that therapists' observer-rated FIS predicted those therapists' real-life clinical outcomes. Subsequent studies confirmed that FIS scores predicted outcome and alliance (Anderson, Crowley, et al., 2016; Anderson, McClintock, et al., 2016) and that these skills were trainable through deliberate practice (Anderson et al., 2020; Rousmaniere et al., 2017). The effects of observer-rated therapist

interpersonal skills on treatment outcomes were similarly demonstrated by Schöttke et al. (2017) in a 5-year longitudinal study with postgraduate students. However, these results depend on the perspective of a trained rater who should have a solid knowledge of psychology and psychotherapy, but whose observations may differ from the client's actual experience. FIS seems to improve and maintain a solid therapeutic alliance (Wampold & Imel, 2015). However, the existing scales that measure these are intended for external observers and not therapy clients, which are the only source of information regarding their own perspectives and feelings about the therapist and therapeutic relationship. To fill this gap, we propose the creation of the FIS-C, which is an adaptation of the original FIS (Anderson et al., 2009) and FIS-IS scales (Uhlin, 2011).

The FIS-C is designed to be applied directly to psychotherapy clients. It's a short and intuitive scale that can be administered after each session without compromising its time and dynamics.

While psychotherapy has proven effective, research suggests that 5%–14% of clients deteriorate while in treatment and that therapists cannot effectively identify such cases (Lambert, 2013). Client feedback has been demonstrated to reduce deterioration and consequently improve therapy outcomes (Lambert & Shimokawa, 2011). Boswell et al. (2015) have highlighted the importance of implementing routine outcome monitoring in clinical practice to identify negative events and clients at risk of deterioration as well as to act in service of positively reverting the change process. We propose the development and validation of the FIS-C scale, which allows the assessment, through the perspective of the client, of specific therapist characteristics that have been demonstrated to influence therapy outcomes. This measure makes it possible for clients to share their views on specific therapist behaviors that can be undermining the therapy process and to express opinions that might be more difficult to express directly to the therapist. In turn, the therapist benefits from accessing important information about their performance and some possible “blind spots.”

The FIS-C has potential implications for clinical practice, through the in-session assessment of fluctuations in each therapy process, and process-outcome research, through the assessment of clients' overall perspective of therapists' performance. This data can, on the one hand, help the therapist practice and improve and, on the other, facilitate the discussion in the therapy of important issues regarding the therapist and the relationship, which may cause ruptures, or even dropouts, if not addressed in time.

Finally, this scale can complement other existing measures, based on self-reports and expert raters, enabling a more comprehensive assessment of the therapist's performance. This might represent new possibilities for process-outcome research on therapist effects and client feedback.

2 | METHODS

2.1 | Participants

The sample for this study was comprised of 194 Portuguese clients aged 18 and older, and attending psychotherapy. They were on average 34.33 years old ($SD = 12.05$), and mostly women ($N = 138$; 71.10%). In terms of education, more than half reported having attended higher education ($N = 120$; 61.90%). Concerning employment status, the majority were employed ($N = 105$; 54.10%), followed by students ($N = 68$; 35.10%). Finally, about a third of the participants reported being in psychotherapy for more than a year ($N = 75$; 38.70%). All participants consented to participate in this study.

2.2 | Instruments

Three distinct instruments were employed for this study.

2.2.1 | Facilitative interpersonal skills-client version

An adaptation of the original FIS (Anderson et al., 2009) and FIS-in session (FIS-IS) scales (Uhlin, 2011) was conducted. A seven-item, 5-point Likert scale was created where each item related to various dimensions of interpersonal skills: verbal fluency, emotional expression, persuasion, acceptance and understanding, hope and positive expectations, empathy, and collaboration. The extremes of the Likert scale were adjusted to fit the client's perspective of the therapist. For example, in the FIS-IS scale, a score of 1 on the verbal fluency item described as *The participant has great difficulty verbalizing his or her ideas throughout the course of the session (e.g., may sound anxious, shaky, or timid throughout the session). The participant lacks confidence in speaking and is consistently difficult to follow* was modified in the FIS-C to *My therapist had great difficulty verbalizing his/her ideas throughout the session (e.g. sounded anxious). He/she lacked confidence in speaking and was consistently difficult to follow.*

The FIS-C was administered to the clients, who were asked to answer based on their last therapy session. The client version of this instrument aims to understand how clients perceive their therapist in terms of their abilities to create and develop a therapeutic relationship and help solve the problems that led that person to therapy.

Two other instruments were deployed, to assess discriminant and convergent validity: one with little conceptual overlap (clinical outcomes in routine evaluation-outcome measures [CORE-OM]-10), and another with a high degree of relatedness (session rating scale [SRS]).

2.2.2 | Clinical outcomes in routine evaluation-outcome measures-10

The CORE-OM (Evans et al., 2002), which was adapted to the Portuguese population (Sales et al., 2012), is a 10-item self-report questionnaire on a 5-point Likert scale which is used as a screening tool to measure outcomes focusing on the most important generic aspects of psychological wellbeing. This scale is formulated in a negative manner; e.g., higher scores are indicative of poorer outcomes.

2.2.3 | Session rating scale

The SRS (Duncan et al., 2003; Miller et al., 2000) is used to evaluate the quality of the therapeutic alliance in psychotherapy. It is comprised of four items on a 10-point Likert scale, where each individual item explores the dimensions of relation, objectives and themes, approach, and generalities.

2.3 | Procedure

A model using the theoretical unidimensional FISC-C structure was specified through a confirmatory factor analysis, and its fit was adjudged. Subsequently, we proceeded with an assessment of the specific psychometric properties of the FIS-C—validity (factorial, discriminant, convergent, and nomological), reliability, and sensitivity. The analyses reported here were conducted using IBM SPSS 27, IBM AMOS 26, and James Gaskin's StatsToolsPackage Macro.

3 | RESULTS

3.1 | Confirmatory factor analysis

In this section, we aimed to test the unidimensional structure of FIS-C. For model estimation, maximum likelihood was used, one of the most commonly used options due to its robustness to potential deviations from normality (Arbuckle, 2007). Model fit was adjudged based on multiple indicators, as dictated by good statistical practices (Hu & Bentler, 1999): for the purposes of this exercise, we used the χ^2 goodness-of-fit test, the corresponding X^2 statistic, its derivate X^2/df index, the comparative fit index—CFI, and the parsimony-adjusted CFI variant—PCFI, the root-mean-square error of approximation—RMSEA. For comparative purposes between models, whenever required, we employed Akaike's information criterion (AIC) and the Browne–Cudeck criterion (BCC).

In this section, two models will be described. The first one refers to the FIS-C scale alone, whereas the second model adds the SRS and CORE-10 scales to adjudge discriminant and convergent validity. Although it would perhaps be more parsimonious to conduct the exercise in a singular model, creating a measurement model containing the FIS-C scale alone allows a better judgment of model fit and factorial loadings on the scale's merits by itself (while the second model also incorporates SRS and CORE-10 into fit computation, which is useful for model evaluation but less so for scale development).

3.2 | Model I—FIS-C measurement model

As previously mentioned, Model I consists of the measurement model for FIS-C by itself. Specification of the exploratory factor analysis(EFA) structure yielded a model with a good fit ($\chi^2(14) = 22.746$, $p = 0.065$; $X^2/df = 1.625$; CFI = 0.991; PCFI = 0.660; RMSEA = 0.057; $p[\text{rmsea} \leq 0.05] = 0.355$). As such, it was not necessary to fall back to exploratory approaches such as evaluating modification indices (Bollen, 2014), and the analysis essentially confirmed the EFA-extracted as well as the theoretical factorial structure. Thus, we were able to proceed to the next stage of validation, which concerns itself with validity, reliability, and sensitivity. Figure 1 illustrates the measurement model for FIS-C.

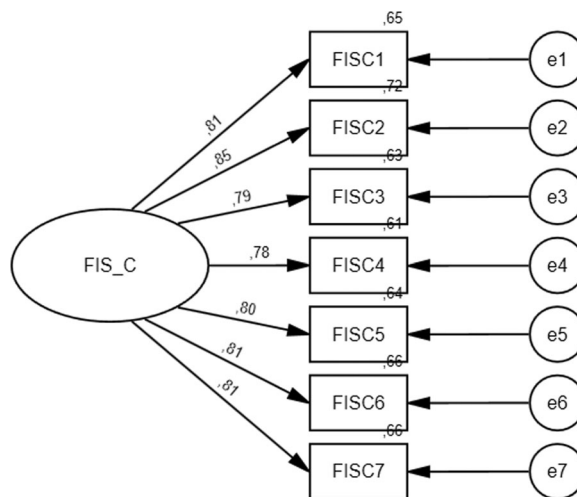


FIGURE 1 Measurement model for FIS-C with standardized regression weights (loadings). Ellipses indicate latent variables, and squares indicate manifest variables. Error terms are indicated by the latent variables labeled “e.” FIS-C, facilitative interpersonal skills-client version

3.3 | Model II—FIS-C, SRS, and CORE-10 measurement model

As previously described, this second model introduced the SRS and CORE-10 scales as additional constructs to evaluate discriminant and convergent validity. Since these are well-established scales, we will not focus on their own psychometric properties and factorial loadings, as they are used merely for contrasting purposes. The fit of this new measurement model was slightly worse than Model I in relative terms, but still deemed as good in absolute terms ($\chi^2(186) = 376.813$, $p < 0.001$; $X^2/df = 2.026$; CFI = 0.929; PCFI = 0.823; RMSEA = 0.073; $p[\text{rmsea} \leq 0.05] < 0.01$). As such, the analysis proceeded with an evaluation of the remaining psychometric properties. To conduct the validity assessment, the StatsToolsPackage macro (Gaskin, 2016) was employed, producing a global validity assessment that is summarized in Table 1, which will be described throughout the remainder of this section.

3.4 | Validity

Four different facets of validity were evaluated in this study: factorial validity, convergent validity, discriminant validity, and nomological validity (Hair et al., 2014).

3.4.1 | Factorial validity

To demonstrate factorial validity, all individual items of the construct must have standardized loadings into the respective latent variable exceeding the threshold of 0.50 (Marôco, 2010). As previously shown during the estimation of Model I (see Table 1), this was attained for all items; as such, factorial validity was confirmed.

3.4.2 | Convergent validity

This aspect requires that all items on average exhibit high loadings into their respective construct, that is, they converge into it. A strict measure of convergent validity can be obtained through the average variance extracted (AVE; Fornell & Larcker, 1981). The Fornell-Larcker criterion establishes that to demonstrate convergent validity, AVE must exceed the threshold of 0.50 (Fornell & Larcker, 1981). FIS-C exhibited an AVE of 0.652, meeting the required criterion, thus confirming its convergent validity.

TABLE 1 Validity and reliability evaluation

	CR	AVE	MSV	ASV	Correlations		
					FIS-C	SRS	CORE
FIS-C	0.929	0.652	0.707	0.478	0.808		
SRS	0.934	0.781	0.707	0.465	0.841	0.883	
CORE-10	0.849	0.402	0.248	0.235	-0.498	-0.472	0.634

Note: The diagonal of the correlation matrix indicates the square root of the AVE.

Abbreviations: ASV, average shared variance; AVE, average variance extracted; CORE, clinical outcomes in routine evaluation; CR, composite reliability; FIS-C, facilitative interpersonal skills-client version; MSV, maximum shared variance; SRS, session rating scale.

3.4.3 | Discriminant validity

Demonstrating discriminant validity requires contrasting the construct with other constructs and demonstrating a low degree of interfactor correlations and cross-loadings. This can be done by calculating the maximum shared variance (MSV), which is the square of the highest inter-factorial correlation, and the average shared variance (ASV), which is the average of the sum of squared interfactorial correlations. The criteria for demonstrating discriminant validity are twofold; first, the square root of the AVE must be greater than all interfactorial correlations; second, the AVE for a given factor must be greater than its MSV and ASV. As shown in Table 2, discriminant validity is confirmed when FIS-C is contrasted with CORE-10 (a scale with negligible conceptual overlap) but is not confirmed when contrasted with SRS (a scale with some degree of conceptual relatedness), due to their very high correlations ($r = 0.841$). This can either be indicative of a threat to face validity (e.g., the formulation of the items makes the client's perception of FIS indistinguishable from his or her perceived session rating), or alternatively as an indication of positive nomological validity (e.g., higher FIS are conducive to a higher session rating) (Hair et al., 2014). As these are two competing and mutually exclusive interpretations, a complementary diagnostic exercise was required to assess which interpretation was correct.

A variant model was specified, in which the SRS latent variable was excluded, and all SRS items were loaded into FIS-C instead. The fit of this variant model was then compared with Model II's fit to assess whether it was significantly different. The rationale behind this procedure is that a significantly worse fit on the variant model provides evidence for the separation of constructs (Hair et al., 2014). This was done through a χ^2 test, contrasting the baseline Model II ($\chi^2(186) = 376.813$) with the variant Model II ($\chi^2(188) = 538.088$). The differences between model fits were significant, ($\chi^2\Delta(2) = 161.275$, $p < 0.001$), and CFIs indicated a worsening of the fit when SRS items were considered as part of the FIS-C construct ($AIC_{\text{variant}} = 624.088$; $BCC_{\text{variant}} = 635.152$; vs. $AIC_{\text{baseline}} = 466.813$; $BCC_{\text{baseline}} = 478.392$), suggesting that these constructs, albeit related, are not one and the same. As such, we argue that discriminant validity is demonstrated.

3.4.4 | Nomological validity

Nomological validity requires that the correlations between constructs are in line with what is expected based on existing theory (Hair et al., 2014). FIS-C exhibited a significant and negative correlation with CORE-10 ($r = -0.498$), indicating that higher levels of FIS are associated with better therapeutic outcomes. This is in line with existing literature (Anderson et al., 2009). As previously discussed, FIS-C is also highly correlated with SRS ($r = 0.841$). The original FIS scale was previously found to be a significant predictor of quality in therapeutic working alliance, which

TABLE 2 Sensitivity evaluation of the FIS-C items

	Min	Max	M	SD	Sk	Ku
FISC1	1	5	4.35	0.893	-1.577	2.708
FISC2	1	5	4.20	0.931	-1.230	1.302
FISC3	1	5	4.16	0.957	-1.089	0.761
FISC4	1	5	4.38	0.787	-1.292	1.727
FISC5	1	5	4.08	0.940	-0.921	0.707
FISC6	1	5	4.24	0.873	-1.190	1.362
FISC7	1	5	4.27	0.865	-1.288	1.904

Abbreviation: FIS-C, facilitative interpersonal skills-client version.

is a component of the SRS scale (Uhlen, 2011). Because we excluded the alternative possibility—as discussed in the previous point of this section—we interpret this correlation as indicative of an expected relation. As such, evidence is also provided for the nomological validity of the FIS-C.

3.5 | Reliability

To assess FIS-C's reliability, we employed the composite reliability (CR) indicator (Fornell & Larcker, 1981), considered a superior alternative to the traditionally used Cronbach's α (Marôco, 2010). The threshold for acceptable reliability is 0.70 (Hair et al., 2014). With a CR of 0.929, FIS-C is demonstrated to be a highly reliable scale.

3.6 | Sensitivity

The last of the psychometric properties to be tested is sensitivity, which relates to the scale's ability to discriminate between two distinct individuals. The criteria for demonstrating this is that each individual item exhibits sufficiently normal distribution (Marôco, 2010). A common criterion for establishing this is when an item's skewness and kurtosis are under the absolute value of 3 (Kline, 2016). As shown in Table 2, this condition was met for all items; thus, the sensitivity of the scale is also demonstrated.

3.7 | Test-retest reliability

A separate sample was collected in a longitudinal manner to compute test-retest reliability and ascertain whether the FIS-C measurements exhibited temporal stability. A sample of 55 clients was recruited, and FIS-C was administered in three separate moments, with 5 weeks between administrations. However, due to experimental mortality, only 31 clients completed the survey in the three moments.

Intraclass correlations were computed for the FIS-C mean across the three moments, yielding a score of 0.703 using average measures, indicating some temporal fluctuation. Observing the FIS-C scores over time reveals how these changes occur, as shown in Table 3.

4 | DISCUSSION

Literature on psychotherapy efficacy reveals that specific therapist characteristics are important for therapeutic success (Wampold et al., 2017; Wampold, 2017). More precisely, the construct of FIS encompasses a set of important therapist qualities that have been shown to have a robust impact on psychotherapy outcomes (Anderson, Crowley, et al., 2016; Anderson, McClintock, et al., 2016).

TABLE 3 FIS-C scores over three time periods

	N	Min	Max	M	SD
T1	55	3.71	5	4.68	0.360
T2	40	4.00	5	4.72	0.321
T3	31	4.43	5	4.85	0.200

Abbreviation: FIS-C, facilitative interpersonal skills-client version.

The development and validation of the FIS-C scale derived from an important question: If part of the reason why some therapists are more effective than others is that they vary on specific interpersonal therapist qualities, why not ask clients to assess these characteristics and use that information to target areas of improvement?

Therapists are usually too optimistic when assessing their clinical competence (Walfish et al., 2012) and differ from clients in the assessment of important variables such as the therapeutic alliance (e.g., Hartmann et al., 2015). Research also shows that client ratings of the therapeutic alliance are better predictors of outcomes than those of therapists (Bachelor & Horvath, 1999).

Moreover, most psychotherapists have great difficulty identifying the existence of clinical deterioration in their clients (Hatfield et al., 2010; MacDonald, 2015).

With this study, we propose a new way to improve the effectiveness of the therapeutic process. Through the information collected with the FIS-C, clinicians will be able to adjust their interventions based on client feedback. As discussed, monitoring clinical progress and using client feedback to tailor psychotherapy interventions to clients has been shown to have a positive impact on therapy outcomes, potentially reducing undesirable therapy effects (Lambert & Shimokawa, 2011). The process of developing and validating the FIS-C scale in this study revealed that the instrument has appropriate psychometric properties, making it a new possible option for clinical use.

In terms of validity, factorial validity and convergent validity were fully demonstrated; for discriminant validity, we argue that it is also demonstrated given the rationale which was described in-depth in the respective section; and finally, evidence is also provided for FIS-C's nomological validity. In terms of reliability, the scale was shown to be highly reliable; in terms of test–retest, the scale shows consistency over time, although some fluctuations do occur. Finally, in terms of sensitivity, all items were demonstrated to have a sufficiently normal distribution, thus providing evidence for the scale's sensitivity.

The strong correlation between this scale and the SRS suggests some interdependency between relational process variables and specific therapist skills. Because the “Person of the Therapist” is considered a common factor that accounts for effects on relational process variables (see American Psychological Association Division 29 Task Force on empirically supported relationships; Norcross & Lambert, 2011), the therapist's specific skills will likely influence these variables.

Some process variables can actually be defined as therapist skills. For instance, empathy can be defined as the therapist's skillfulness in being empathic and conveying understanding to their client (Anderson, McClintock, et al., 2016). The authors argue that it is not yet clear whether there are numerous independent therapist skills or if there is a more generic interpersonal relatedness factor. They assume, however, that therapy relationships that include effective common factors like empathy, warmth, and persuasiveness are also corresponding to a set of skills that differentiate the therapists' ability to be effective and promote an effective therapy process (Anderson & Patterson, 2013). Despite that, we argue that it is possible for a therapist to engage in effective therapy processes without a clear understanding of which particular skills are enabling them. Therefore, even though for process-outcome research purposes there are similarities in terms of construct definition between the FIS and other process variables, when studying therapist effects and expertise, this scale proves more useful as it isolates specific trainable skills.

The growing importance of the therapist as a cross-theoretical variable requires a reformulation of clinical training based on clinical expertise and expert performance. We argue that the FIS-C scale can become an important tool in the recognition of therapist blind spots—such as lack of empathy and difficulty resolving alliance ruptures—informing the clinician of potential problems in the therapy process of a particular client or across clients. Interpersonal skills can be practiced and improved to increase effectiveness. The amount of time spent deliberately training specific therapeutic skills has proven effective in increasing clinical performance and has been demonstrated to be a significant predictor of client outcomes (Miller et al., 2013). Therefore, the implementation of client feedback instruments such as the one we sought to validate in this study allows the therapist to understand which skills are important to train for a specific client and/or their practice in general. This is a promising start to prevent the occurrence of harmful events and premature termination in psychological support services.

5 | LIMITATIONS AND FUTURE DIRECTIONS

Despite the establishment of validity and reliability for the FIS-C scale, the sample consisted solely of Portuguese nationals. This means that the instrument is valid and reliable for populations that are culturally similar and should not be generalized to contrasting populations or cultures. An appropriate use of this measure should include the cultural aspects inherent to the client's background.

Client feedback alone, however, is insufficient to assess therapists' performance in an accurate and comprehensive way. Both clients' and therapists' perceptions, characteristics, and behaviors can influence the clients' view of their therapists' FIS. Responsiveness might pose an "issue" to research on therapist competence as clients' perspectives might be biased by the context and their own needs (Stiles, 2021). Nonetheless, this scale might be an important addition to other measures of therapist performance as it is important to gather multiple perspectives to get a valid assessment.

Future research on the FIS-C should focus on its validation for other populations and/or countries. Moreover, we suggest that future studies with the FIS-C focus on longitudinal assessment of therapist skills and therapist improvement based on client feedback. It would also be interesting to develop a longitudinal study where the FIS-C and FIS-in session scales can be both used to assess discrepancies or correlations between client-rated and observer-rated therapist areas of improvement. The importance of using this scale in clinical practice will likely be reinforced through further studies.

6 | CONCLUSION

The literature on FIS is based on the establishment of the therapist as an important variable for therapy outcomes. Unlike other factors that account for outcome variance, the variables associated with therapists and their skills are more easily identifiable and manageable, through the application of feedback instruments. The FIS-C was validated in this study and represents a new possible resource for therapists to understand their clients' points of view on the skills that can be practiced and improved. It is suggested that the complementary use of this measure with other existing ones is the most comprehensive way to measure therapist performance. Moreover, the combination of client feedback and deliberate practice of client-targeted skills can ultimately contribute to higher therapist effectiveness and outcome improvement.

CONFLICT OF INTEREST

The authors declare no conflict of interest.

DATA AVAILABILITY STATEMENT

The data that support the findings of this study are available on request from the corresponding author. The data are not publicly available due to privacy or ethical restrictions.

ORCID

João M. Santos  <http://orcid.org/0000-0001-6414-6891>

Matilde Barata  <http://orcid.org/0000-0002-0521-2171>

Sara Rathenau  <http://orcid.org/0000-0001-5067-8029>

Inês Amaro  <http://orcid.org/0000-0002-1119-6059>

Alexandre Vaz  <http://orcid.org/0000-0001-8443-0064>

Daniel Sousa  <http://orcid.org/0000-0001-8213-4616>

PEER REVIEW

The peer review history for this article is available at <https://publons.com/publon/10.1002/jclp.23469>.

REFERENCES

- Anderson, T., Crowley, M. E. J., Himawan, L., Holmberg, J. K., & Uhlin, B. D. (2016). Therapist facilitative interpersonal skills and training status: A randomized clinical trial on alliance and outcome. *Psychotherapy Research, 26*(5), 511–529. <https://doi.org/10.1080/10503307.2015.1049671>
- 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.
- 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. (2013). *Facilitative interpersonal skill task and rating method* (PhD thesis). Ohio University.
- Anderson, T., Perlman, M. R., McCarrick, S. M., & McClintock, A. S. (2020). Modeling therapist responses with structured practice enhances facilitative interpersonal skills. *Journal of Clinical Psychology, 76*(4), 659–675. <https://doi.org/10.1002/jclp.22911>
- Arbuckle, J. (2007). *Amos 16.0 user's guide*. SPSS Inc.
- Bachelor, A., & Horvath, A. (1999). The therapeutic relationship, *The heart and soul of change: What works in therapy* (pp. 133–178). American Psychological Association.
- 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.). John Wiley & Sons.
- Beutler, L., Malik, M., Alimohamed, S., Harwood, T., Talebi, H., Noble, S., & Wong, E. (2004). In M. J. Lambert, A. E. Bergin, & S. L. Garfield (Eds.), *Therapist variables* (5th ed.). Wiley.
- Bollen, K. A. (2014). *Structural equations with latent variables*. John Wiley & Sons.
- Bordin, E. S. (1979). The generalizability of the psychoanalytic concept of the working alliance. *Psychotherapy: Theory, Research & Practice, 16*, 252–260.
- Boswell, J. F., Kraus, D. R., Miller, S. D., & Lambert, M. J. (2015). Implementing routine outcome monitoring in clinical practice: Benefits, challenges, and solutions. *Psychotherapy Research, 25*(1), 6–19. <https://doi.org/10.1080/10503307.2013.817696>
- Castonguay, L. G. & Hill, C. E. (Eds.). (2017). *How and why are some therapists better than others? Understanding therapist effects*. American Psychological Association.
- Duncan, B. L., Miller, S. D., Sparks, J. A., Claud, D. A., Reynolds, L. R., Brown, J., & Johnson, L. D. (2003). The session rating scale: Preliminary psychometric properties of a “working” alliance measure. *Journal of Brief Therapy, 3*(1), 3–12.
- Fornell, C., & Larcker, D. F. (1981). Evaluating structural equation models with unobservable variables and measurement error. *Journal of Marketing Research, 18*, 39–50.
- Frank, J. D., & Frank, J. B. (1993). *Persuasion and healing: A comparative study of psychotherapy* (3rd ed.). Johns Hopkins University Press.
- Gaskin, J. (2016). Stats Tool Package. <http://statwiki.kolobkreations.com>
- Hair, J. F., Black, W. C., Babin, B. J., & Anderson, R. E. (2014). *Multivariate data analysis*. Pearson Education Limited.
- Hartmann, A., Joos, A., Orlinsky, D. E., & Zeeck, A. (2015). Accuracy of therapist perceptions of patients' alliance: Exploring the divergence. *Psychotherapy Research, 25*(4), 408–419. <https://doi.org/10.1080/10503307.2014.927601>
- Hatfield, D., McCullough, L., Frantz, S. H., & Krieger, K. (2010). Do we know when our clients get worse? An investigation of therapists' ability to detect negative client change. *Clinical Psychology & Psychotherapy, 17*(1), 25–32. <https://doi.org/10.1002/cpp.656>
- Hu, L., & Bentler, P. M. (1999). Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. *Structural Equation Modeling: A Multidisciplinary Journal, 6*(1), 1–55.
- Kline, R. B. (2016). *Principles and practice of structural equation modeling*. Guilford Press.
- Lambert, M. J. (Ed.). (2013). *Bergin and Garfield's handbook of psychotherapy and behavior change* (6th ed.). John Wiley & Sons.
- Lambert, M. J., & Shimokawa, K. (2011). Collecting client feedback. *Psychotherapy, 48*(1), 72–79.
- MacDonald, N. E. (2015). Vaccine hesitancy: Definition, scope and determinants. *Vaccine, 33*(34), 4161–4164. <https://doi.org/10.1016/j.vaccine.2015.04.036>
- Marôco, J. (2010). *Análise de equações estruturais: Fundamentos teóricos, software & aplicações*. ReportNumber, Lda.
- Miller, S., Duncan, B., & Johnson, L. (2000). *The session rating scale 3.0*. Springer.
- Miller, S. D., Hubble, M. A., Chow, D. L., & Seidel, J. A. (2013). The outcome of psychotherapy: Yesterday, today, and tomorrow. *Psychotherapy, 50*(1), 88–97. <https://doi.org/10.1037/a0031097>

- Norcross, J. C., & Lambert, M. J. (2011). *Psychotherapy relationships that work II* (Vol. 48, No. 1, p. 4). Educational Publishing Foundation.
- Norcross, J. C., & Lambert, M. J. (2019). *Psychotherapy relationships that work: Evidence-based therapist contributions* (Vol. 1). Oxford University Press.
- Norcross, J. C., & Wampold, B. E. (Eds.). (2019). *Psychotherapy relationships that work: Evidence-based therapist responsiveness* (Vol. 2). Oxford University Press.
- Rousmaniere, T., Goodyear, R. K., Miller, S. D. & Wampold, B. E. (Eds.). (2017). *The cycle of excellence: Using deliberate practice to improve supervision and training*. John Wiley & Sons.
- Sales, C. M. D., Moleiro, C. M. M., Evans, C., & Alves, P. C. G. (2012). Versão portuguesa do CORE-OM: Tradução, adaptação e estudo preliminar das suas propriedades psicométricas. *Archives of Clinical Psychiatry*, 39(2), 54–59.
- Schöttke, H., Flückiger, C., Goldberg, S. B., Eversmann, J., & Lange, J. (2017). Predicting psychotherapy outcome based on therapist interpersonal skills: A five-year longitudinal study of a therapist assessment protocol. *Psychotherapy Research*, 27(6), 642–652. <https://doi.org/10.1080/10503307.2015.1125546>
- Stiles, W. B. (2021). Responsiveness in psychotherapy research: Problems and ways forward. In J. C. Watson & H. Wiseman (Eds.), *The responsive psychotherapist: Attuning to clients in the moment* (pp. 15–35). American Psychological Association. <https://doi.org/10.1037/0000240-002>
- Uhlir, B. (2011). Therapist in-session rated facilitative interpersonal skills (FIS-IS) in the psychotherapy process. Ohio University.
- Walfish, S., McAlister, B., O'Donnell, P., & Lambert, M. J. (2012). An investigation of self-assessment bias in mental health providers. *Psychological Reports*, 110(2), 639–644. <https://doi.org/10.2466/02.07.17.PRO.110.2.639-644>
- Wampold, B. E. (2017). What Should We Practice? 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* (pp. 49–65). John Wiley & Sons.
- Wampold, B. E., Baldwin, S. A., & Imel, Z. E. (2017). *What characterizes effective therapists?* American Psychological Association.
- Wampold, B. E., & Imel, Z. E. (2015). *The great psychotherapy debate: The evidence for what makes psychotherapy work* (2nd ed.). Routledge.

How to cite this article: Santos, J. M., Barata, M., Rathenau, S., Amaro, I., Vaz, A., Sousa, D., Severino, M., & Taveira, M. (2022). Development and validation of the facilitative interpersonal skills scale for clients. *Journal of Clinical Psychology*, 1–12. <https://doi.org/10.1002/jclp.23469>