Determining differences between therapists using an extended version of the facilitative interpersonal skills performance test

Thiel, S.J. van; Jong, K. de; Misset, K.S.; Joosen, M.C.W.; Klink, J.J.L. van der; Vermunt, J.K.; Dam, A. van

Received: 29 May 2023 | Accepted: 20 March 2024

DOI: 10.1002/jclp.23687

RESEARCH ARTICLE

WILEY

Determining differences between therapists using an extended version of the facilitative interpersonal skills performance test

Sabine J. van Thiel[1,2] | Kim de Jong[3] | Kirsten S. Misset[3] | Margot C. W. Joosen[1] | Jac J. L. van der Klink[1,4] | Jeroen K. Vermunt[1] | Arno van Dam[1,2]

[1]Tilburg School of Social and Behavioral Sciences, Tranzo–Academic Collaborative Center Work & Health, Tilburg University, Tilburg, The Netherlands

[2]Mental Health Institute GGZ WNB, Bergen op Zoom, The Netherlands

[3]Institute of Psychology, Leiden University, Leiden, The Netherlands

[4]Optentia, North West University of South Africa, Potchefstroom, South Africa

Correspondence

Sabine J. van Thiel, Tranzo, Tilburg University, Professor Cobbenhagenlaan 125, 5037 DB Tilburg, The Netherlands.

Email: s.j.vanthiel@tilburguniversity.edu

Abstract

Objectives: The therapist-facilitative interpersonal skills (FIS) has shown to predict therapy outcomes, demonstrating that high FIS therapists are more effective than low FIS therapists. There is a need for more insight into the variability in strengths and weaknesses in therapist skills. This study investigates whether a revised and extended FIS-scoring leads to more differentiation in measuring therapists' interpersonal skills. Furthermore, we explorative examine whether subgroups of therapists can be distinguished in terms of differences in their interpersonal responses.

Method: Using secondary data analysis, 93 therapists were exposed to seven FIS-clips. Responses of therapists using the original and the extended FIS scoring were rated.

Results: Three factors were found on the extended FIS scoring distinguishing supportive, expressive, and persuasive interpersonal responses of therapists. A latent profile analysis enlightened the presence of six subgroups of therapists.

Conclusion: Using the revised and extended FIS-scoring contributes to our understanding of the role of interpersonal

skills in the therapeutic setting by unraveling the question what works for whom.

KEYWORDS
common factors, interpersonal skills, performance-based assessment, therapist effects

1 | INTRODUCTION

The interpersonal skills of therapists appear to play an important role in explaining differences in therapists' average effectiveness. Research shows that some therapists have consistently more successful treatments than other therapists (Johns et al., 2019). For example, Saxon and Barkham (2012) found recovery rates between 20% and 96%, with below-average effective therapists having twice as many cases of deterioration as highly effective therapists. Therapists' interpersonal skills and performance-based tests continue to be mainly predictive of therapists effectiveness (Anderson et al., 2009, 2016; Heinonen & Nissen-Lie, 2020; Lingiardi et al., 2018; Schöttke et al., 2017). In particular, interpersonal skills demonstrated in response to a difficult patient or challenging situation are important to predict psychotherapy outcomes (Anderson et al., 2009; Schöttke et al., 2017).

The facilitative interpersonal skills (FIS) performance task was designed to "elicit responses that are indicants of a person's ability to perceive, understand, and communicate a wide range of interpersonal messages, as well as a person's ability to persuade others with personal problems to apply suggested solutions to their problems and abandon maladaptive patterns" (Anderson et al., 2009). Anderson et al. (2019) rests the construction and definition of the FIS on a collection of identified relationship factors as researched and described by Norcross and Lambert (2018). Norcross and Lambert (2011, 2018) describe empirically supported therapist behaviors that improve treatment outcomes, such as empathy, congruence, emotional expression, and (repairing rupture) alliances, and conclude that psychotherapy relationships make substantial and consistent contributions to outcomes independent of the type of treatment. Anderson et al. (2019) assumes that these relationship factors can be translated as therapist level skills. The FIS performance task was therefore developed in an attempt to identify therapist factors with an eye toward controlling this factor in the experimental study.

In the FIS test, therapist respond to a client video as if they were their therapist (Anderson et al., 2009). The therapist's responses are recorded by video and scored afterward on eight items, namely verbal fluency, hope and positive expectations, persuasiveness, emotional expression, warmth, acceptance, & understanding, empathy, alliance bond capacity, and alliance rupture-repair responsiveness. Anderson et al. (2009) presented a video of a challenging patient to 25 therapists at a college counseling center. In a multilevel model, FIS scores at the therapist level were used to predict the improvement of 1141 patients seen by the therapists. The FIS was a strong predictor of patient improvement, with a correlation of 0.47 between FIS and outcome. In a similar study, FIS scores has predicted treatment outcomes of patients who were seen for eight or fewer sessions with high FIS therapist having significant better outcomes than low FIS therapists (Anderson et al., 2016).

The power, impact, and originality of the FIS instrument has inspired many researchers, proven by the many replications and follow-up studies that have been established in the past decade (Bate & Tsakas, 2022; Goldberg et al., 2020; Munder et al., 2019; Santos et al., 2022; Steggles & de Jong, 2018; van Thiel et al., 2021; Zech et al., 2022). The FIS was translated into eight different non-English languages, and additional versions of the FIS clips are released focusing on different difficult therapy situations, such as cultural differences and suicidal situations (Anderson et al., 2019). A text based FIS was developed and validated (Zech et al., 2022). Also a FIS client version was developed and validated to collect clients' perspectives on relevant therapist characteristics (Santos et al., 2022). Research also examined therapists' physiological reactions during the presentation of provocative and

benign FIS clips (Steggles & de Jong, 2018). Bate and Tsakas (2022) developed a FIS task using child and adolescent patients as stimuli. Other research has focused on evaluating whether machine learning tools can automatically assess therapist interpersonal skills (Goldberg et al., 2020). The FIS instrument had made a major contribution to research as a predictor of outcomes and therapist differences.

Although the FIS makes a great contribution to understanding differences between therapists, there are also a number of limitations with regard to the instrument. The first limitation is that, with only one factor, the instrument only distinguishes "high" versus "low" FIS therapists. Several authors argue that therapists adapt their interpersonal style to the state of interaction (Hatcher, 2015; Lambert, 2013; Lambert & Ogles, 2014; Stiles, 2013; Swift & Derthick, 2013). Literature on appropriate responsiveness argue that a therapist is effective because "he or she consistently does the right thing, which may be different each time, providing each client with a different, tailored treatment" (Stiles & Horvath, 2017). There is a need to study what types of therapists are effective for what types of patients (Krause & Lutz, 2009). Yet, the interaction hypothesis in which some therapists express what types of interpersonal behavior with some types of patients, which then produces better outcomes, cannot be further explored with the FIS instrument at this point.

The second limitation lies in the fact that some FIS scoring items have a double representation. Certain skills were measured as a condition for another skill to have a high score and were therefore not evaluated independently. This causes item-construct validity problems, and in the end, it remains unknown what the (un) successful element was in predicting outcomes.

And finally, as a result of not measuring certain constructs separately, the weight in the scoring has largely shifted to measuring therapist's ability of creating an emotional bond but very much less on skills that bring a client to change. Balancing both support and change is an important but challenging interpersonal mission for therapists. Confronting difficult truths can cause the therapist to encounter resistance (Wachtel, 1993). On the other hand, research has shown that when therapists are too accommodating, they find it difficult to bring a patient to change. Therapist friendly-submissiveness was significantly negatively associated with patient-rated depth, alliance, and overall helpfulness (Regas et al., 2017). A clear measurement of FIS of therapists focusing on changing the maladaptive behavior of a client is lacking.

For this study, the FIS-scoring was revised and extended with three items, namely: agency, reframing, and congruence. The item "agency" assesses the extent to which the therapist actively manages the client's own responsibility and state of being in action and does not respond to the appeal of the client to take over power, save the other person, or to make the therapist feel guilty or inferior. The item "reframing" assesses the extent to which the therapist says something that puts the client's situation in a different context or perspective. The item "congruence" measures the therapist's capacity to transparency give his/her voice to the other person and not hiding behind a professional role or holding back feelings that are obvious in the encounter. Congruence (or genuineness) was first described by Carl Rogers (1957). A meta-analysis of Kolden et al. (2018) indicates that better psychotherapy outcomes, for example, reduced symptoms, increased functioning, and well-being, can be expected when the therapist is seen as congruent/genuine.

The general aim of this study was to explorative investigate in what way a revised and extended FIS scoring would lead to more differentiation in therapists' interpersonal skills. Bordin (1979) described the therapeutic alliance as a function of both developing a strong affective bond but also having agreement on desired outcomes and methods used. We hypothesize that a revised and extended FIS scoring leads to a two-dimensional structure representing both affect and change orientations. Furthermore, we explorative examine whether subgroups of therapists can be distinguished in terms of differences in their interpersonal responses. Finally, we analyze whether therapist variables predict interpersonal responses. We assume that therapists variables do not predict interpersonal style since therapists' age, gender, clinical experience, theoretical orientation, and professional degree are not consistently linked in research to therapist effectiveness (Bjaastad et al., 2018; Huppert et al., 2001; Owen et al., 2016; Wampold & Brown, 2005).

2 | METHOD

2.1 | Data

The study is a secondary analysis of data in which 93 Dutch native speaking cognitive behavioral therapists from two large mental healthcare organizations participated. The sample consists of cognitive behavioral therapy trained psychologists. All participants have obtained at least a master of science title in clinical psychology, whether or not supplemented with a post-master's training as a healthcare psychologist or psychotherapist. Age, years of experience, gender, and nationality were surveyed as demographic variables.

2.2 | Procedure

The therapists were recruited on-site and informed about the purpose, aim, and procedure of the study through presentations given by a member of the research group. There was also information provided on paper, and a (online) video was recorded with information that could be viewed privately or when the participant missed the face-to-face presentation. Therapists participated from different branches throughout the Netherlands. The therapist received a gift certificate of 10 euro's and were allowed to do the test in worktime. The research assistants prepared and started the task but left the room when the therapists started. This was done at the therapists' workplace in a quiet room where the therapist would not be disturbed. Therapists received a prerecorded instruction about the task. Therapists were exposed to seven Dutch FIS clips (van Thiel et al., 2021). Before each stimulus clip, the therapists received a brief explanation about the client and his/her problems. As briefly discussed above, therapists were then asked to respond as if they were the clients therapist and in a way they think is helpful (see Anderson et al., 2009, 2016, 2018).

2.3 | Task

For this study, the seven original Dutch language FIS clips were used. The psychometric properties of the FIS clips was tested in earlier research. Good content validity and reasonable interrater reliability was indicated (van Thiel et al., 2021). The clips are distributed from a variety of interpersonal situations, leaving the participant with some form of request or appeal (Anderson et al., 2018). This means that the clips bring different interpersonal challenges and that they are differently positioned on the interpersonal circumplex space of control and affiliation (Kiesler & Schmidt, 2006). Each clip shows a different client (with different interpersonal challenges), with the names "John," "Bonnie," "Less," "Suzie," "Lauren," "Hillary," and "Jack." The clips were also presented in this fixed order.

2.4 | Original FIS scoring

In accordance with the translation of the Dutch FIS clips (van Thiel et al., 2021), the FIS scoring was translated by two independent translators. A synthesis measurement took place and the back translation method was used to assure linguistic equivalence. The original FIS scoring describes eight FIS items: "verbal fluency," "hope and positive expectations," "persuasiveness," "emotional expression," "warmth, acceptance, and understanding," "empathy," "alliance bond capacity," and "alliance rupture-repair responsiveness." Each item is rated on a 5-point Likert scale ranging from 1 (*not characteristic*) to 5 (*extremely characteristic*). The manual with a description of the original FIS scoring can be found or requested via https://www.fisresearch.com.

2.5 | Extended FIS scoring

The extended FIS scoring is generally the same as the original scoring and adds newly formulated items "agency," "reframing," and "congruence." These items were also rated similarly to the original scoring on a 5-point Likert scale ranging from 1 (*not characteristic*) to 5 (*extremely characteristic*).

Agency was also described in the original soring in item 2, "hope and positive expectations." We removed this condition from item 2 and described it in a newly formulated item, "agency." Agency measures the extent to which the participant continues to actively manage the client's own responsibility and does not respond to the appeal of the client. To receive a 5-point score, the participant discusses the appeal that's being made to him/her and also discusses the clients own responsibility. For a 1-point score, the participant fully responds to the appeal and appears to feel fully responsible for the clients behavior.

Similarly, in the original scoring, to receive a high score on item 3, "persuasiveness," the participants "must offer some rationale or reframing of the other's experience." We also removed this condition from item 3 and described it in a newly formulated item, "reframing." Reframing measures the extend in which the participant would say something that puts the clients situation in a different context or perspective and give a different view on the situations with the aim of causing an inner, interpersonal, cognitive, or behavioral change (Castonguay & Hill, 2012). To receive a 5-point score, the participant's response is very contradictory and confrontation to the clients perspective. Whereas a 1-point score, the participant avoids any form of reflection or reformulation of the clients perspective.

To help scoring, examples were written down of possible responses that would receive a high score on agency and reframing for each FIS-clip. To illustrate, this is the example that was given to the raters to help score agency and reframing of the clip of Jessica: "*Jessica's appeal is that the therapists gets the feeling that he needs to apologize. An example in which the appeal is not answered is when the therapists returns to Jessica that it is her decision whether she continues to come to the therapy session. Or even; if she does something in which she doesn't see a point of doing, she shouldn't do it.*"

Congruence measures the participant's ability to communicate his or her experience with the client honestly and in a conscientious manner (Kolden et al., 2011). It refers to explicitly discussing one's feelings, wants, and needs; the participant says what he/she feels and thinks. An insincere, indirect style, on the other hand, refers to verbal messages that camouflage and conceal the true intentions, wants, or needs. For a 5-point score, the participant indicates very explicitly what his feelings and/or thoughts are, whereas for a 1-point score, the participants says (manifest level) is completely different from his/her nonverbal (latent level) communication.

2.6 | Raters

Therapists' responses to the FIS clips were scored on both the original and the extended FIS scoring in two separate settings. This had to do with another study for which the original scoring was used. Two clinical master students of Leiden University and Tilburg University rated all responses of therapists using the extended FIS scoring, including de revised FIS items. Two other clinical master students of Leiden University rated the responses of the therapists using the original FIS scoring. Both settings had separate calibration sessions in which therapist responses or ambiguities in scoring options were discussed, for example, cases in which therapists gave a terse reaction. In the beginning, these meetings were held weekly. As the rates got more routine, the frequency of the sessions changed to once every 2 week and further on once every 4 weeks. All responses were coded separately. Aggregated scores were used for the analyses.

2.7 | Statistical analysis

All statistical analyses were conducted using IBM SPSS statistics version 26, using a two-tailed significance level of $p < .05$. The mean scores of each FIS item for every therapist were calculated and used for the analyses except for the interrater reliability. Interrater reliability was measured using the intraclass correlation coefficient (ICC). The interpretation of ICC values, according to Cicchetti (1994), is as follows: less than 0.40 indicates poor reliability, valuables between 0.40 and 0.59 are indicated as fair, values of 0.60 and 0.74 are indicated as good, and values of 0.75 and 1.00 as excellent reliability. A principal component analysis with oblimin rotation and maximum likelihood on both the original FIS scoring and then the extended FIS scoring was performed.

To find clusters of therapists, we performed a latent class analysis using the 11 FIS items as indicators. These indicators were treated as continuous, and normally distributed, and uncorrelated within classes, implying the analysis can also referred to as latent class cluster analysis or latent profile analyses. Estimation was performed using LatentGOLD, version 6.0 (Vermunt & Magidson, 2021). Two of the program's default settings were changed: (1) because of the rather small sample size, we used a model with equal residual variances between classes (instead of unequal ones), and (2) to avoid local maxima, the number of start sets and iterations per set were increased to 320 and 250, respectively. The decision on the number of clusters was mainly based on the BIC (Bayesian information criterion), but we also looked the AIC (Akaike information criterion), the total and the maximum of bivariate residual (BVR), and the Vuong–Lo–Mendell–Rubin (VLMR) test.

After deciding on the number of clusters, we obtained the classifications. These were subsequently used to investigate the relationship between cluster membership and age and experience. For this purpose, we used the Step3-Dependent option with the proportional class assignment and BCH adjustment for classification error in LatentGOLD (see Bolck et al., 2004; Vermunt, 2010).

3 | RESULTS

3.1 | Preliminary analyses and interrater reliability

Out of all the therapists' ($N = 89$) responses to the clips ($N = 596$), 27 clip-responses (4.3%) were missing due to technical problems. The sample age range was 23–70 (M = 33.94, SD = 8.57) years. Eight participants identified as male (9%) and 77 (86.5%) as female. Four (4.5%) participants were labeled as unknown because they did not fill in their gender. The years of experience ranged between less than 1 year and 35 years (M = 7.31, SD = 8.57). Nationality was largely Dutch. Seven participants indicated that they had a different background, mainly Surinamese (3), Afghan (1), Turkish (1), and Asian (1).

The total mean FIS score of the original FIS scoring was 3.38, ranging from 2.65 to 4.16. The total mean FIS score of the extended FIS scoring was 3.42, ranging from 2.61 to 4.53. Table 1 shows descriptive statistics of the extended FIS scoring for each FIS domain. ICC were calculated for each item (see Table 1). All FIS domains got good reliability except for verbal fluency, hope and positive expectations, persuasiveness, and alliance-bond capacity with excellent reliability.

3.2 | Principal component analysis

Analyzing the original FIS scoring, we found one factor explaining 70.3% of the variance, with factor loadings ranging from 0.76 for verbal fluency to 0.90 for persuasiveness. We found three factors on the extended FIS scoring explaining 80.4% of the variance (see Table 2). Factor 1 explained 52.8% of the variance and was comprised of six items: "hope and positive expectations," "warmth acceptance and understanding," "empathy," "alliance bond

TABLE 1 Descriptive statistics and intraclass correlation coefficient (ICC).

FIS item	M (SD)	Intraclass correlation coefficient (ICC) (95% confidence interval)	*F* test with true value 0			
			Value	*df*1	*df*2	*p* Value
1. Verbal fluency	3.37 (0.62)	0.78 (0.73–0.82)	4.67	592	592	<.001
2. Hope and positive expectations	3.20 (0.30)	0.77 (0.72–0.81)	4.56	592	592	<.001
3. Persuasiveness	3.44 (0.55)	0.78 (0.74–0.81)	4.51	592	592	<.001
4. Emotional expression	3.43 (0.53)	0.74 (0.69–0.78)	3.82	592	592	<.001
5. Warmth, acceptance, and understanding	3.64 (0.50)	0.73 (0.68–0.77)	3.67	592	592	<.001
6. Empathy	3.35 (0.46)	0.66 (0.58–0.71)	3.02	592	592	<.001
7. Alliance bond capacity	3.36 (0.43)	0.80 (0.77–0.83)	5.18	592	592	<.001
8. Alliance rupture-repair responsiveness	3.33 (0.58)	0.74 (0.64–0.81)	4.30	592	592	<.001
9. Reframing	3.15 (0.47)	0.75 (0.65–0.81)	4.40	592	592	<.001
10. Agency	3.48 (0.46)	0.72 (0.67–0.76)	3.59	592	592	<.001
11. Congruence	3.85 (0.51)	0.69 (0.57–0.78)	3.63	590	590	<.001

Abbreviations: df, degrees of freedom; FIS, facilitative interpersonal skills.

Average measures, two-way mixed effect model, and absolute agreement.

capacity," "alliance rupture repair responsiveness," and "congruence" factor. This factor seems to measure therapists' ability to affirm, validate, and sympathize with the client. We labeled this the supportive-response dimension. Factor 2 explained 15.07% of the variance and was comprised of four items: "verbal fluency," "persuasiveness," "emotional expression," and "congruence." This dimension seems to measure therapists' ability to express themselves in an affective and powerful way. It is not so much about what is said but much more about how it is said. We labeled this dimension the persuasive-response dimension. Factor 3 explained 12.48% of the variance and was comprised of three items: "alliance-rupture repair responsiveness," "reframing," and "agency." This dimensions seems to measure therapists' abilities to initiate change and development. We labeled this the expressive-response dimension.

3.3 | Therapist differences in their interpersonal response style

Table 3 reports the relevant statistics for the estimated models with one to seven latent classes. As can be seen, based on the BIC, one would select the six-class solution. The AIC and VLMR tests indicate a model with seven or

TABLE 2 Principal component analysis on the extended FIS scoring.

FIS domain	Factor 1	Factor 2	Factor 3
1. Verbal fluency	−0.10	**1.02**	−0.12
2. Hope and positive expectations	**0.65**	0.23	−0.06
3. Persuasiveness	0.10	**0.83**	0.18
4. Emotional expression	0.23	**0.57**	0.12
5. Warmth, acceptance, and understanding	**0.85**	0.11	−0.04
6. Empathy	**0.79**	0.15	0.14
7. Alliance bond capacity	**0.87**	0.16	−0.07
8. Alliance rupture-repair responsiveness	**0.57**	−0.04	**0.57**
9. Reframing	−0.03	−0.01	**0.95**
10. Agency	−0.10	0.04	**0.83**
11. Congruence	**0.33**	**0.32**	0.25

Note: Factor 1, supportive-response dimension; Factor 2, persuasive-response dimension; and Factor 3, expressive-response dimension.

Abbreviation: FIS, facilitative interpersonal skills.

even more classes should be selected. Comparison of the total and maximum BVR values with the ones of the one-class model indicates that overall the dependencies between the 11 indicators are picked up rather well by the six-class model, though the maximum BVR indicates that some item pairs still have a somewhat larger residual dependency. The total and maximum BVR do not improve by increasing the number of clusters. Since the inclusion of free covariances for these pairs did not change the encountered six-cluster solution, we decided to keep the standard six-cluster model as our final model. As can be seen from the large entropy R^2 value, class membership can be predicted very well using the persons' scores on the 11 indicators.

It should be noted that the selected six-cluster model contained four larger classes and two very small classes, each consisting of only two therapists. Almost identical larger classes we found in the four-class and the five-class model, and one of the small classes was also found in five-class model.

In the six-cluster model (see Figure 1), the first class was composed of 36.7% of the sample ($n = 33$) and represented therapists with an average score on all items (we will refer to this cluster as the average-profile [A-profile]). The second class was composed of 26.0% of the sample ($n = 23$) and represented therapists with an average score but with less focus on warmth and empathy, they make a smooth chat and actively manage the clients responsibility for action (named average-expressive-persuasive-profile [AEP-profile]). The third class, composed of 17.9% of the sample ($n = 16$), was characterized by therapists who had a slightly higher average score, in particular, they are more empathic and congruent but focus less on confronting the client (named high-average-supportive-persuasive-profile [HASP-profile]). The fourth class, composed of 14.5% of the sample ($n = 13$), was characterized by therapists who perform slightly less across the board, there is a focus on the therapeutic relationship, being kind and understanding toward the client, but in doing that avoid conflict and confrontation, and are not expressive or convincing (named low-average-supportive-profile [LAS-profile]). The fifth class, composed of 2.4% of the sample ($n = 2$) and represented therapists who have high scores on all items; are warm, empathic, expressive, persuasive, congruent, and have great rupture-repair abilities (named high-supportive-expressive-persuasive-profile [HSEP-profile]). Finally, the last class, composed of 2.4% of the sample ($n = 2$) and represented therapists who are less warm and not focusing on building and maintaining a good therapeutic bond but strongly focus on managing the clients own responsibility for action and change (named low-expressive-profile [LE-profile]).

TABLE 3 Cluster proportions, cluster-specific means, significance test for differences in means, and explained variances.

	Cluster 1	Cluster 2	Cluster 3	Cluster 4	Cluster 5	Cluster 6	Wald	*p* Value	R^2
Cluster proportions	0.37	0.26	0.18	0,.14	0.02	0.02			
HPE	3.27	3.04	3.49	2.98	3.83	2.40	128.70	<.001	0.60
WAU	3.75	3.32	4.26	3.21	4.63	2.40	263.88	<.001	0.76
EMP	3.41	3.03	3.96	2.91	4.72	2.81	356.22	<.001	0.82
ABC	3.48	3.04	3.70	3.23	4.36	2.18	122.31	<.001	0.60
VF	3.41	3.37	3.79	2.64	4.24	3.30	41.04	<.001	0.33
PER	3.48	3.42	3.91	2.62	4.63	3.15	107.28	<.001	0.58
EE	3.45	3.35	3.87	2.71	4.76	3.90	108.49	<.001	0.57
REFR	3.04	3.26	3.41	2.69	4.19	3.42	42.31	<.001	0.35
AG	3.32	3.69	3,.73	3.03	4.44	3.75	48.81	<.001	0.40
ARRR	3.30	3.22	3.98	2.73	4.87	2.26	128.02	<.001	0.63
CONG	3.93	3.69	4.31	3.11	4.84	4.38	122.42	<.001	0.59
Age	35.79	31.41	32.89	34.37	32.50	39.50	13.33	0.02	0.05
Experience	8.74	5.96	6.74	6.79	6.75	12.00	30.94	<.001	0.04

Abbreviations: ABC, alliance bond capacity; AG, agency; ARRR, alliance rupture repair responsiveness; CONG, congruence; EE, emotional expression; EMP, empathy; HPE, hope and positive expectations; PER, persuasiveness; REFR, reframing; VF, verbal fluency; WA, warmth, acceptance, and understanding.

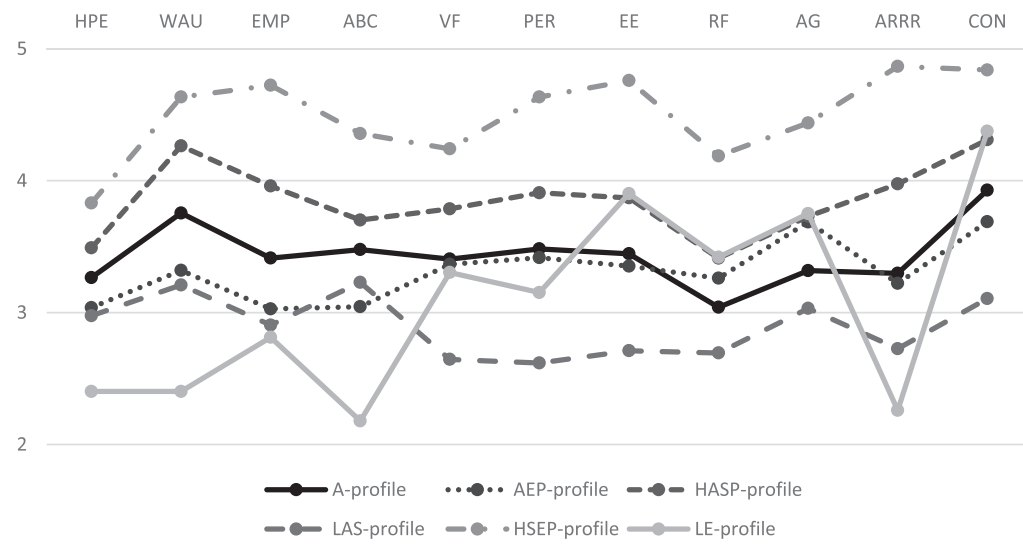

FIGURE 1 Profile graph for 6-cluster model. The A-profile, average profile; AEP-profile, average-expressive-persuasive-profile; HASP-profile, high-average-supportive-persuasive-profile; HSEP-profile, high-supportive-expressive-persuasive-profile; LAS-profile, low-average-supportive-profile; LE-profile, low-expressive-profile.

A pairwise Wald test for age and years of experience between cluster 6 and any other cluster was statistically significant. Showing that therapists in cluster 6 are older and more experienced compared with the other clusters. The other five clusters did not differ significantly on age or years of experience.

4 | DISCUSSION

A good therapist must be prepared to support and confront patients to make them reflect themselves. The FIS instrument provides insight into the interpersonal skills of therapists. The aim of this study was to investigate whether a revised and extended FIS scoring, including items measures of agency, reframing, and congruence, would lead to more differentiation in measuring therapists' interpersonal skills and behavior.

The results showed one factor for the original FIS scoring suggesting being an unidimensional scale. These results are consistent with previous studies (Anderson et al., 2009, 2016; Munder et al., 2019). Based on the literature, we expected to find two dimensions, namely supportive and change, for the revised and extended FIS scoring. This expectation was reflected in the results, but on top of that, we found a third dimension. The persuasive dimension is different from the other dimension in the principle that it is not focusing on what the therapists says but really on how it is said. Our study does not clarify how the underlying items relate to each other. Specifically for the items "congruence" and "alliance rupture repair responsiveness," it remains unclear how these construct can exactly be understood within the interpersonal skills dimensions. Possibly these construct can be seen as products of other skills. For example, emotional, warmth, empathic, and fluent responses give the observer the impression of the therapist being honest and being his/her genuine self. The importance of congruence has been much discussed in scientific literature yet only been measured using questionnaires (Kolden et al., 2018).

An explorative examination of whether subgroups of therapist can be distinguished in term of differences in their interpersonal responses with the latent cluster analyses indicated the presence of six clusters: (1) the A-profile, (2) the AEP-profile, (3) the HASP-profile, (4) the LAS-profile, (5) the HSEP-profile, and (6) the LE-profile. The LE-profile appeared to be significantly represented by older and more experienced therapists. This confirms that age and experience are not necessarily predictors of success, in this case even the opposite. The representation of a large average group with smaller outliers at the top and bottom fits the patterns of therapists variability found in the study of Saxon and Barkham (2012). Yet, to our knowledge, this is the first study investigating therapist differences using latent cluster analyses in which we get a first impression of how exactly "supershrinks" (Miller et al., 2008; Okiishi et al., 2003) or underperformers differ from other therapists in terms of their interpersonal skills. Whereas high performers (HSEP-profile) both invest in a strong affective relationship on which confrontation can be dealt with, it seems that underperformers (LE-profile) lack to invest in a strong affective bond but strongly focus on managing the clients own responsibility for action and change. Giving the appearance of a therapist that says something to stop the client being lazy; causing him/her to act. In the middle area, differences between therapists, ranging from passive types (LAS-profile) to more action-oriented therapists (AEP-profile), are found.

The question is what the effect of these different styles is on the outcome of the treatment. It could well be that clients differ in the approach they need to improve. Some clients may benefit more from a therapist who is warm and empathetic, and others may benefit more from a therapist who points out their own responsibility in the change process.

The revised and extended scoring makes it possible to distinguish therapists by their interpersonal skills. To prove the value of replacing/adding the items in terms of predictive validity, more research is necessary. Further research with this scoring could lead to a better understanding whether different response styles are more effective for which type of patient. The hypothesis can be based on the interpersonal theory (Kiesler, 1996) and the attachment theory (Bowlby, 1988), emphasizing that opposite or contrasting interpersonal orientations in the therapist and client are optimal for the process and the outcome of the treatment relationship. In FIS research, this would mean matching change-oriented therapists with passive, ineffective clients and matching dominant, aggressive clients with affective-oriented therapists could indeed improve outcomes. Answers can also be sought within aspects of the theory of proximal development. The

zone of proximal development indicates differences between what somebody can learn without guidance and what a person can achieve with guidance. In the translation to psychotherapy, this means that a therapist has to attune to the client; setting the bar for change not too high but also challenge a person to expend perceptions and skills.

4.1 | Limitations and recommendations

The following limitations can have implications for further research and applicability. It is possible that our findings were influenced by our homogeneous sample of female, protocolled trained cognitive behavioral therapists. Our sample consisted almost entirely of Dutch white women. Therefore we do not know whether the results are generalizable to therapists from other cultural or ethnic groups. Additionally, our sample was relatively small. As a result, two subgroups, such as the LE-profile, consisted of only two therapists. This means that results were sensitive to outliers. A larger and more diverse sample of therapists, for example, from different countries, with different theoretical backgrounds, could provide more clarity about the generalizability of the results to other settings, countries, or cultures.

High interrater reliabilities were found. This could say something about the strength of the different dimensions found. On the other hand, it can say something about a rating bias about the criteria. Furthermore, because of the explorative nature of the study, no pre-registration was done. In a subsequent study, one might consider using more experienced therapists in assessing responses. Pre-registration, more raters, and/or using a more detailed description of narratives could reduce the risk of outcome bias.

5 | CONCLUSION

Revising the FIS scoring, adding three additional items (agency, reframing, and congruence) leads to differentiation in measuring therapist interpersonal skills. Six subgroups of therapists can be distinguished in terms of their differences in interpersonal responses. Research employing the revised FIS scoring will contribute to extending our understanding of the role of interpersonal skills in the therapeutic setting. The extended FIS scoring is directly applicable for educational purposes. Using the extended rating in further research, limitations must be taken into account.

ACKNOWLEDGMENTS

The authors would like to thank HSK Group and all respondents who participated in the study. This work was supported by GGZ Westelijk Noord-Brabant (GGZ WNB).

CONFLICT OF INTEREST STATEMENT

The authors declare no conflict of interest.

DATA AVAILABILITY STATEMENT

Permission to reproduce data was granted. Derived data supporting the findings of this study are available from the corresponding author upon reasonable request.

ETHICS STATEMENT

All participating therapists provided written informed consent in which they agreed to participate in the study and agreed with using their data for further analyses, scientific publications, and presentations. All participants in this study were given the opportunity to (re−) consider participation and submit questions before accepting the informed consent. Participation did not entail any compensation. All data was treated in an aggregated and anonymous way. This study was reviewed and approved by the Psychology Research Ethics Committee at Leiden University.

ORCID

Sabine J. van Thiel http://orcid.org/0000-0003-2403-2229

Kim de Jong http://orcid.org/0000-0002-7621-9290

PEER REVIEW

The peer review history for this article is available at https://www.webofscience.com/api/gateway/wos/peer-review/10.1002/jclp.23687.

REFERENCES

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.

Anderson, T., Patterson, C., McClintock, A. S., McCarrick, S. M., & Song, X., The Psychotherapy and Interpersonal Lab Team. (2018). *Facilitative interpersonal skills task and rating manual. Unpublished rating manual*. Ohio University.

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.

Bate, J., & Tsakas, A. (2022). Facilitative interpersonal skills are relevant in child therapy too, so why don't we measure them? *Research in Psychotherapy: Psychopathology, Process and Outcome*, *25*(1), 595.

Bjaastad, F. J., Wergeland, H. G. J., Haugland, M. B. S., Gjestad, R., Havik, O. E., Heiervang, E. R., & Öst, L. G. (2018). Do clinical experience, formal cognitive behavioural therapy training, adherence, and competence predict outcome in cognitive behavioural therapy for anxiety disorders in youth? *Clinical Psychology & Psychotherapy*, *25*(6), 865–877.

Bolck, A., Croon, M., & Hagenaars, J. (2004). Estimating latent structure models with categorical variables: One-step versus three-step estimators. *Political Analysis*, *12*(1), 3–27.

Bordin, E. S. (1979). The generalizability of the psychoanalytic concept of the working alliance. *Psychotherapy: Theory, Research & Practice*, *16*(3), 252.

Bowlby, J. (1988). *A secure base: Parent-child attachment and healthy human development*. Basic Books.

Castonguay, L. G., & Hill, C. E. (2012). *Transformation in psychotherapy: Corrective experiences across cognitive behavioral, humanistic, and psychodynamic approaches*. American Psychological Association.

Cicchetti, D. V. (1994). Guidelines, criteria, and rules of thumb for evaluating normed and standardized assessment instruments in psychology. *Psychological assessment*, *6*(4), 284.

Goldberg, S. B., Flemotomos, N., Martinez, V. R., Tanana, M. J., Kuo, P. B., Pace, B. T., Villatte, J. L., Georgiou, P. G., Van Epps, J., Imel, Z. E., Narayanan, S. S., & Atkins, D. C. (2020). Machine learning and natural language processing in psychotherapy research: Alliance as example use case. *Journal of Counseling Psychology*, *67*(4), 438–448.

Hatcher, R. L. (2015). Interpersonal competencies: Responsiveness, technique, and training in psychotherapy. *American Psychologist*, *70*(8), 747–757.

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

Huppert, J. D., Bufka, L. F., Barlow, D. H., Gorman, J. M., Shear, M. K., & Woods, S. W. (2001). Therapists, therapist variables, and cognitive-behavioral therapy outcome in a multicenter trial for panic disorder. *Journal of Consulting and Clinical Psychology*, *69*(5), 747–755.

Johns, R. G., Barkham, M., Kellett, S., & Saxon, D. (2019). A systematic review of therapist effects: A critical narrative update and refinement to review. *Clinical Psychology Review*, *67*, 78–93.

Kiesler, D. J. (1996). *Contemporary interpersonal theory and research: Personality, psychopathology, and psychotherapy*. John Wiley & Sons.

Kiesler, D. J., & Schmidt, J. A. (2006). *The impact message inventory-circumplex (IMI-C) manual: Sampler set, manual, test booklet, scoring key, work sheets*. Mind Garden.

Kolden, G. G., Klein, M. H., Wang, C. C., & Austin, S. B. (2011). Congruence/genuineness. *Psychotherapy*, *48*(1), 65–71.

Kolden, G. G., Wang, C. C., Austin, S. B., Chang, Y., & Klein, M. H. (2018). Congruence/genuineness: A meta-analysis. *Psychotherapy*, *55*(4), 424–433.

Krause, M. S., & Lutz, W. (2009). Process transforms inputs to determine outcomes: Therapists are responsible for managing process. *Clinical Psychology: Science and Practice*, *16*(1), 73–81.

Lambert, M. J. (2013). The efficacy and effectiveness of psychotherapy. Chapter 6. In M. J. Lambert (Ed.), *Bergin and Garfield's handbook of psychotherapy and behavior change* (6th ed., pp. 169–218). Wiley.

Lambert, M. J., & Ogles, B. M. (2014). Common factors: Post hoc explanation or empirically based therapy approach? *Psychotherapy*, *51*, 500–504.

Lingiardi, V., Muzi, L., Tanzilli, A., & Carone, N. (2018). Do therapists' subjective variables impact on psychodynamic psychotherapy outcomes? A systematic literature review. *Clinical Psychology & Psychotherapy*, *25*(1), 85–101.

Miller, S. D., Hubble, M., & Duncan, B. (2008). Supershrinks: What is the secret of their success. *Psychotherapy in Australia*, *14*(4), 14–22.

Munder, T., Schlipfenbacher, C., Toussaint, K., Warmuth, M., Anderson, T., & Gumz, A. (2019). Facilitative interpersonal skills performance test: Psychometric analysis of a German language version. *Journal of Clinical Psychology*, *75*(12), 2273–2283.

Norcross, J. C., & Lambert, M. J. (2011). *Psychotherapy relationships that work II* (Vol. 48, p. 4). Educational Publishing Foundation.

Norcross, J. C., & Lambert, M. J. (2018). Psychotherapy relationships that work III. *Psychotherapy*, *55*(4), 303–315.

Okiishi, J., Lambert, M. J., Nielsen, S. L., & Ogles, B. M. (2003). Waiting for supershrink: An empirical analysis of therapist effects. *Clinical Psychology & Psychotherapy: An International Journal of Theory & Practice*, *10*(6), 361–373.

Owen, J., Wampold, B. E., Kopta, M., Rousmaniere, T., & Miller, S. D. (2016). As good as it gets? Therapy outcomes of trainees over time. *Journal of Counseling Psychology*, *63*(1), 12–19.

Regas, S. J., Kostick, K. M., Bakaly, J. W., & Doonan, R. L. (2017). Including the self-of-the-therapist in clinical training. *Couple and Family Psychology: Research and Practice*, *6*(1), 18–31.

Rogers, C. R. (1957). The necessary and sufficient conditions of therapeutic personality change. *Journal of Consulting Psychology*, *21*, 95–103.

Santos, J. M., Barata, M., Rathenau, S., Amaro, I., Vaz, A., Sousa, D., Severino M., & Teixeira, M. (2022). Development and validation of the facilitative interpersonal skills scale for clients. *Journal of Clinical Psychology*, *79*(4), 1166–1177.

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.

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.

Steggles, K., & de Jong, K. (2018). *Does therapist emotion regulation moderate their facilitative interpersonal skills? Paper presented at the 49th International Society for Psychotherapy Research International Annual Meeting, Amsterdam, The Netherlands.*

Stiles, W. B. (2013). The variables problem and progress in psychotherapy research. *Psychotherapy*, *50*, 33–41.

Stiles, W. B., & Horvath, A. O. (2017). *Appropriate responsiveness as a contribution to therapist effects*.

Swift, J. K., & Derthick, A. O. (2013). Increasing hope by addressing clients' outcome expectations. *Psychotherapy*, *50*, 284–287.

van Thiel, S., Joosen, M. C. W., Joki, A. L., van Dam, A., van der Klink, J. J. L., & de Jong, K. (2021). Psychometric analysis of the Dutch language facilitative interpersonal skills (FIS) video clips. *Research in Psychotherapy: Psychopathology, Process, and Outcome*, *24*(1), 513.

Vermunt, J. K. (2010). Latent class modeling with covariates: Two improved three-step approaches. *Political Analysis*, *18*, 450–469.

Vermunt, J. K., & Magidson, J. (2021). *Upgrade manual for latent GOLD basic, advanced/syntax, and choice version 6.0*. Statistical Innovations Inc.

Wachtel, P. L. (1993). *Therapeutic communication: Principles and effective practice*. Guilford Press.

Wampold, B. E., & Brown, G. S. (2005). Estimating variability in outcomes attributable to therapists: A naturalistic study of outcomes in managed care. *Journal of Consulting and Clinical Psychology*, *73*(5), 914–923.

Zech, J., Foley, V. K., Hull, T. D., & Anderson, T. (2022). Assessing the quality of digital patient-therapist communication: The development and validation of a text-based facilitative interpersonal skills task. *Psychotherapy Research*, *33*(6), 743–756.

How to cite this article: van Thiel, S. J., de Jong, K., Misset, K. S., Joosen, M. C. W., van der Klink, J. J. L., Vermunt, J. K., & van Dam, A. (2024). Determining differences between therapists using an extended version of the facilitative interpersonal skills performance test. *Journal of Clinical Psychology*, *80*, 1698–1710. https://doi.org/10.1002/jclp.23687