Therapist Belief Scale: Empirical validation in a sample of Spanish psychologists and university students

Rocío Fernández-Velasco[1], Carmen Casares-Guillén[1], Isabel Silva-Lorente[1], Douglas J. Boegaerts[1], Paula Moya-García[1] and Helena Garrido-Hernansaiz[1,2]

[1] *Centro Universitario Cardenal Cisneros, Universidad de Alcalá, Madrid, Spain*

[2] *Universidad Pontificia de Comillas, Madrid, Spain*

Short title: Therapist Belief Scale

Corresponding author: Carmen Casares Guillén, Centro Universitario Cardenal Cisneros. Av. Jesuitas, 34, 28806 Alcalá de Henares, Madrid. España.
E-mail: carmen.casares@cardenalcisneros.es

Título abreviado: Escala de Creencias del Terapeuta

ABSTRACT

Objective: The aim of this study was to translate and validate the Therapist Belief Scale (TBS) into Spanish, and instrument designed to examine therapists' negative beliefs on therapy and care quality. **Method:** A total of 513 participants, including students and psychology professionals, were involved in assessing the scale's reliability and validity. Factorial validity was examined through confirmatory factor analyses, while criterion-related validity was evaluated by correlating TBS scores with measures of anxiety sensitivity, trait anxiety, optimism, and self-efficacy. **Results:** We found a bifactor structure based on 13 items, comprising three specific factors (low tolerance for distress, inflexibility, and responsibility-control) and a general factor. The TBS demonstrated good reliability and validity, showing positive correlations with anxiety sensitivity

and trait anxiety and negative correlations with optimism and self-efficacy. Sensitivity analyses further supported its discriminative capacity. **Conclusions**: The TBS is a reliable and valid tool for assessing therapists' beliefs, offering valuable insights for future research on psychologists' beliefs about professional performance, training, and supervision.

Keywords: Therapist beliefs; TBS; anxiety; optimism; self-efficacy.

Escala de Creencias del Terapeuta: Validación empírica en psicólogos y estudiantes universitarios españoles

RESUMEN

Objetivo: El presente estudio tuvo como finalidad traducir y validar al español la *Therapist Belief Scale* (TBS), escala diseñada para evaluar las creencias negativas de los terapeutas sobre la intervención psicológica. **Método:** Participaron 513 estudiantes y profesionales de la psicología. Se examinó la fiabilidad, la validez factorial (mediante análisis factoriales confirmatorios) y la validez de criterio (mediante las correlaciones de las puntuaciones de la TBS con sensibilidad a la ansiedad, ansiedad rasgo, optimismo y autoeficacia). **Resultados:** Se obtuvo una estructura bifactorial de 13 ítems, constituida por tres factores específicos (baja tolerancia al malestar, inflexibilidad y responsabilidad-control) y un factor general. La TBS mostró buena fiabilidad y validez, correlacionando positivamente con sensibilidad a la ansiedad y ansiedad rasgo, y negativamente con optimismo y autoeficacia. Los análisis de sensibilidad confirmaron su capacidad discriminativa. **Conclusiones:** La TBS demuestra ser una herramienta fiable y válida para medir creencias del terapeuta y su relación con el desempeño profesional, la formación y la supervisión.

Palabras clave: Creencias del terapeuta; TBS; ansiedad; optimismo; autoeficacia.

Introduction

Research on psychotherapy traditionally focused on identifying the most effective treatment for each disorder, but recent studies also emphasize other factors relevant to treatment success, including therapists' characteristics (Castonguay & Hill, 2017; Wampold & Imel, 2015). In their review, Johns et al. (2019) confirmed that therapists indeed make an important contribution to the variability in therapy outcomes. Specifically, professional attributes of therapists showed the strongest relationship to client outcomes (Heinonen & Nissen-Lie, 2019). Tabullo et al. (2023) studied the individual traits and skills that a therapist brings to clinical practice, defined as the "personal style of the therapist", and pointed out that variables such as empathy, orientation, involvement, and adaptability can vary depending on the therapist's theoretical approach, clinical experience, and other personal and contextual variables. The therapist's theoretical approach influences the therapy style and perceived competence (Casari et al., 2019). For instance, psychoanalysts tend to be more spontaneous and emotionally distant; cognitive-behavioral therapists prefer structured interventions and show greater commitment; systemic therapists show lower levels of commitment; and integrative therapists are more spontaneous than cognitive-behavioral therapists (Tabullo et al., 2023). Flexibility in the therapist's personal style, which

includes theoretical approach and technique, facilitates adaptation to the patient's needs. This adaptability improves the therapeutic alliance and allows the therapist to respond effectively to the difficulties of the therapeutic process, promoting better outcomes for the patient (Corbella et al., 2009).

The study of therapist effects thus provides a better understanding of therapist variability and can help improve the effectiveness of therapy (Johns et al., 2019). One core area is how effectively therapists feel able to work toward attaining their goals with patients (Heinonen & Nissen-Lie, 2019). For example, Heinonen et al. (2012) showed that therapists' lower confidence in their effectiveness and skills predicted poorer outcomes, especially in brief treatments. There is also research that analyzes therapists' self-competence; for example, Brosan et al. (2008) found that therapists significantly overrated their competence, particularly those less competent. This finding has serious implications for ensuring effective practice of cognitive therapy in routine clinical situations.

Therapists' beliefs (TB) and attitudes, interacting with their knowledge base, shape emotions and behaviors relevant in therapy selection and implementation (Waller & Turner, 2016). TB can also impact the professional's psychological adjustment, correlating with higher emotional exhaustion and lower personal accomplishment (McLean et al., 2003; Emery et al., 2009). However, despite their significance, TB remain understudied.

The limited studies on TB have often used unpublished or generic scales assessing rigidity (Emery et al., 2009). The Therapist Belief Scale (TBS; Emery et al., 2009; McLean et al., 2003) is the only tool designed to assess a broad set of therapist performance beliefs. However, the TBS is available solely in English, so expanding it to more languages is crucial for assessing psychologists' and psychology students' beliefs in different countries. The information provided by these TB assessments is necessary to help psychologists and students comprehend their effects on therapy (Haarhoff & Kazantzis, 2007) and ensure care quality. This study aimed to conduct a back-translation of the TBS into Spanish and examine the evidence for the validity of the scores among psychologists and psychology students in Spain. By developing a Spanish TB measure, research with professionals and students in Spain will be possible, and the information provided by this measure can directly inform the training and practice of psychotherapy.

The TBS was developed in Australia through a four-step process involving 116 professionals (McLean et al., 2003). From a pool of 58 items, an exploratory factor analysis (EFA) using varimax rotation yielded a final 40-item, one-factor solution, demonstrating an internal consistency of .92. The total score correlated with higher therapist burnout, proving to be a stronger predictor than other work-related variables. However, the method used to perform the EFA was not specified and, as the authors noted, the study was underpowered. A second study (Emery et al., 2009) involved 190 psychologists and considered the 36 items with loadings > .40 in the prior study. Using that 36-item pool, the authors conducted an EFA using principal component analysis (PCA) with a varimax rotation, retaining 29 items with loadings > .40. Four factors emerged: low tolerance of distress (Distress), rigid adherence to the therapeutic model (Inflexibility), responsibility for the outcome (Responsibility), and need for control in therapy and understanding (Control). Then, a confirmatory factor analysis (CFA) was performed keeping the items with loadings > .60, which resulted in 12 items forming three subscales: Distress, Inflexibility, and Control. Reliability was .78 for the total score, .81 for Distress, but lower (.69 and .61) for Inflexibility and Control, indicating reliability concerns in the latter two subscales. Regarding the EFA performed in this second study (Emery et al., 2009), using PCA poses some issues since, although widely used as such, from a methodological point of view it is not a valid method for EFA

(Alavi et al., 2020). Moreover, a varimax rotation is unsuitable for multidimensional constructs like TB, where dimensions are expected to be correlated.

In the present study the validity of the TBS scores will be assessed by examining the associations between TB and various factors. Starting with demographics, gender, and age are potential variables influencing TB. Artkoski and Saarnio (2013) found that female therapists had a more positive attitude toward patients than male therapists. Studies on healthcare providers (Carney et al., 2004, 2007; McCulloch et al., 2005) noted higher uncertainty tolerance in females, linked to TB factors like distress tolerance and need for control and understanding. In addition, female therapists showed greater flexibility, adaptability, and spontaneity, while men adopted more structured styles in their interventions (Tabullo et al., 2023). In this direction, it seems that women focus more on affective empathy, are more detail-oriented, and plan sessions with a more collaborative approach, while men adopt more directive styles and exhibit greater cognitive empathy (Casari et al., 2019).

Regarding age, younger therapists exhibit more distress beliefs in therapy; for instance, Emery et al. (2009) linked younger age with higher endorsement of need for control. Other studies considered that age and years of clinical experience do not seem to be determining variables in the therapist's personal style, although both factors can influence the patient's perception, skill development, confidence, and therapist competence (Tabullo et al., 2023).

Nevertheless, age often correlates with education level and practical experience. Indeed, training influences therapist cognitions (Rameswari et al., 2021). Novice therapists often grapple with anxiety and insecurity regarding their clinical skills, developing their professional identity around feelings of incompetence (Thériault, 2009), and gaining confidence as training advances (Hill et al., 2007). Conversely, experienced professionals hold fewer rigid therapy beliefs (Altabef et al., 2017). As training progresses and experience in therapy grows, negative self-efficacy beliefs diminish (Green et al., 2017; Morrison & Lent, 2018), evident in differences between first- and fourth-year students (Falanga et al., 2014). Hence, varying education and experience levels would likely yield differences in TB, with lower scores among professionals versus students, graduate versus undergraduate students, and fourth-year versus first-year students.

Moreover, training quality (Waller & Turner, 2016), encompassing practical engagement and self-reflection (i.e., supervision and peer interaction; Bennett-Levy & Finley-Jones, 2018; Chigwedere et. al, 2021; Haarhoff et al., 2015; Haarhoff & Kazantzis, 2007), significantly increases students' self-awareness applicable in clinical settings (Lent et al., 2009; Swift et al., 2018). Also, receiving formal training in an area where postgraduate students felt less prepared and more doubtful led to less concern and greater confidence (Tyron, 2001). Consequently, this study expected students involved in targeted TB training during their studies or undergoing learning experiences providing therapist-like involvement and self-reflection (e.g., supervised internships, therapist-simulation project) to exhibit lower scores on the TBS. In Spain, psychology students usually undertake an internship in their fourth and last year, those students are expected to show lower scores on the TBS (Falanga et al., 2014).

The study program modality is a less-explored factor relevant for understanding TB. Abrams et al. (2017) found that students in a mixed online and on-campus counseling master's program scored lower on the counselor exams' Helping Relations section compared to fully on-campus students. On-campus education might offer more practical experience potentially enhancing students' grasp of therapist-client interactions and reducing their TB.

Finally, concerning TB and psychological variables, research reveals negative links with self-efficacy (Lent et al., 2009; Emery et al., 2009) and optimism (Hoy et al., 2008). Additionally, rigid adherence to therapeutic models (a TBS dimension) correlates with heightened anxiety

sensitivity (D'Souza Walsh et al., 2019; Simpson-Southward et al., 2018), and low tolerance for uncertainty (need for control in therapy in the TBS) is linked to increased anxiety (Roscoe & Taylor, 2023). Thus, positive relationships between TBS scores and trait anxiety and anxiety sensitivity are expected, as well as negative associations with self-efficacy and optimism.

In short, this research aimed to translate and validate the TBS into Spanish, a reliable and valid tool for measuring TB. It examines the factorial structure of its scores through CFAs, their internal consistency, and their construct validity by exploring their relationship with related constructs: optimism, trait anxiety, anxiety sensitivity, and self-efficacy. Additionally, it appraises the scores' sensitivity to detect differences in the TB scores based on gender, age, occupation (student, graduate, working), program modality (on-campus, partially on-campus, online), internships, TB training, and work experience. Lastly, it assesses the sensitivity to detect TB changes following a simulated therapist-client experience in undergraduate students.

Methods

Participants

The sample comprised 513 participants who were between 19 and 62 years old, with an average of 25.23 years ($SD = 8.18$). The sample was predominantly female (78.8%). Most participants were college psychology students (86.6%) or practicing clinical psychologist (13.4%, including those in psychology residencies in the National Health System). Among psychology students, 63.6% were undergraduates, while 22.4% pursued master's or PhD programs. Additional sample details are available in Table 1. Furthermore, for some sensitivity analyses, an ancillary sample of 97 third-year psychology students (72.2% women, 25.8% men, 2.1% non-binary; age $M = 21.96$, $SD = 1.02$) was gathered.

Table 1. Sociodemographic characteristics of participants, TBS means, and mean difference tests

Characteristic	n	%	Responsibility-Control					Inflexibility					Distress					Total				
			M	SD	TS*	p	ES	M	SD	TS*	p	ES	M	SD	TS*	p	ES	M	SD	TS*	p	ES
Gender					-1.15	.25	.13			-2.44	.02	.27			-3.15	.002	.35			-2.91	.004	.32
Male	104	20.3	15.57	4.05				10.50	3.63				11.63	4.25				37.69	8.64			
Female	405	78.9	15.04	4.24				9.62	3.18				10.28	3.79				34.94	8.62			
Non-binary [3]	4	0.8	15.75	3.77				10.00	5.89				10.50	4.04				36.25	12.84			
Current occupation					12.48	<.001	.07			7.28	<.001	.04			84.41[2]	<.001	.17			29.30	<.001	.15
Undergraduate student	327	63.6	15.99 [a]	4.18				10.28 [a]	3.06				11.75 [a]	4.00				38.02 [a]	8.39			
Master's degree student	117	22.8	13.90 [b]	3.66				9.03 [b]	3.33				8.71 [b]	2.65				31.6 [b]	6.91			
PhD student [3]	1	0.2	9					4					4					17				
Psychology residency	14	2.7	13.93 [ab]	3.99				10.21 [ab]	3.38				9.14 [b]	2.28				33.29 [ab]	7.41			
Practicing psychologist	55	10.7	13.33 [b]	4.15				8.58 [b]	4.00				7.91 [b]	2.96				29.82 [b]	8.20			
Program modality					2.20	.11	.01			1.78	.17	.01			3.99	.02	.02			4.07	.02	.02
On-campus	373	72.6	15.11	4.17				9.69	3.28				10.32 [a]	3.93				35.11 [a]	8.51			
Partially on-campus	98	19.1	15.78	4.28				10.36	3.34				11.56 [b]	3.66				37.69 [b]	9.12			
Online	43	8.4	14.21	4.18				9.51	3.43				10.42 [ab]	4.14				34.14 [ab]	8.79			
Completed an internship					6.02	<.001	.54			3.55	<.001	.32			10.08[1]	<.001	1.04			8.85[1]	<.001	.83
No	306	59.5	16.05	4.24				10.22	3.07				11.82	3.99				38.09	8.46			
Yes	208	40.5	13.85	3.78				9.18	3.55				8.72	2.97				31.75	7.62			
Any professional experience as a psychologist					4.78	<.001	.50			3.41[1]	<.001	.34			7.67[1]	<.001	.84			6.99	<.001	.73
No	393	76.5	15.64	4.03				10.10	3.10				11.20	3.89				36.94	8.12			
Yes	121	23.5	13.60	4.36				8.82	3.77				8.45	3.26				30.89	8.96			
Received TB training					1.67	.10	.19			1.82	.07	.20			3.87	<.001	.43			3.24	.001	.36
No	414	80.5	15.31	4.13				9.93	3.22				10.89	3.88				36.13	8.30			
Yes	100	19.5	14.53	4.45				9.26	3.61				9.22	3.82				33.01	9.86			

	M	SD	Range	Responsibility-Control		Inflexibility		Distress		Total	
				r	p	r	p	r	p	r	p
Age	25.23	8.18	19-62	-.12	.005	-.16	<.001	-.24	<.001	-.23	<.001

Note. M = Mean. SD = Standard Deviation. TS = Test Statistic (it can be *t, F,* or *H*). ES = Effect size (it can be *d*, δ, or η²). * In homoscedasticity conditions, mean differences between categories of dichotomous variables were assessed via *t*-test and Cohen's *d* effect size statistic was obtained (interpretation: negligible < .20 < small < .50 < medium < .80 < large). For multiple-category variables, one-way ANOVAs were used with post-hoc Tukey tests, assessing the effect size via η² (interpretation: negligible < .01 < small < .06 < medium < .14 < large). Categories with a different superscript letter show a significant difference between them in the psychological impact variable mean.

[1] Homoscedasticity couldn't be assumed, thus the *t*-test results were adjusted for non-homogeneous variances and Glass' δ effect size statistic was used (same interpretation as Cohen's *d*).
[2] Homoscedasticity could not be assumed, thus Kruskal-Wallis' test was used, with *post-hoc* Games-Howell tests.
[3] This category was excluded from the mean comparison analysis because of the small number of participants.

Instruments

    Therapist Belief Scale (TBS; Emery et al., 2009; McLean et al., 2003). It measures TB regarding the effectiveness of professional practice with 29 items on a Likert-type scale. Responses range from 1 (Strongly disagree) to 6 (Strongly agree). The overall Cronbach's alpha for the original version scores was .92 (Mclean et al., 2003). A later 12-item version of the scale (Emery et al., 2009) with three subscales showed an overall internal consistency of .78 (Distress, α = .81; Inflexibility, α = .69; Control, α = .61). See Appendix 1.

    Anxiety Sensitivity Index-3 (ASI-3; Taylor et al., 2007). The Spanish adaptation by Sandín et al. (2007) was used in this study. This 18-item self-report scale assesses physical, cognitive, and social anxiety sensitivity. Item ratings range from high (4) to low (1) agreement levels. The Spanish adaptation reported good internal consistency (α = .91 total score, .84 physical subscale, .87 cognitive subscale, .83 social subscale; Sandín et al., 2007). In this study, only the 5 items from the social dimension were used (α = .79).

    State-Trait Anxiety Inventory – Short Version (STAI; van Knippenberg et al., 1990). A Spanish short form of the STAI was used in this study (Guillén-Riquelme & Buela-Casal, 2013). It assesses trait and state anxiety with 12 items rated on a Likert-type scale (0 to 3). In this study, only the 6-item trait scale, measuring the tendency to perceive stimuli as anxiogenic, was used. Acceptable reliability was observed in the Spanish adaptation (trait anxiety, α = .69) and in this study (α = .68).

    Life Orientation Test-Revised (LOT-R; Scheier et al., 1994). The Spanish adaptation by Ferrando et al. (2002) was used in this study. This measure of generalized optimism comprises 10 items rated on a Likert-type scale ranging from 1 (Strongly disagree) to 5 (Strongly agree). Cano-García et al. (2015) reported adequate internal consistency (α = .73), consistent with this study's findings (α = .80).

    Wallston's Perceived Competence Scale (PCS; Wallston, 1992). The Spanish adaptation by Fernández Castro et al. (1998) was used in this study. This tool assesses self-efficacy with 8 items rated on a Likert-type scale ranging from 1 (Strongly disagree) to 6 (Strongly agree). Cronbach's alpha was .80 in the Spanish adaptation and .86 in this study.

Procedure

    The 29 TBS items (Emery et al., 2009) were back-translated into Spanish. A native bilingual psychologist specializing in research methods translated the items into Spanish and another psychologist fluent in English back-translated them, and small discrepancies were resolved through consensus. The final translation appears in the Appendix.

    Ethics approval was obtained from the first author's university, and participants gave informed consent. Recruitment occurred via institutional emails and various social media platforms during the 2021-2022 academic year and the data were collected through an online questionnaire.

    For the ancillary sample involving the therapist-client simulation, the data were collected in April 2023. Participants were third-year undergraduates who completed the TBS 10 days before and after a simulated therapy session. This training experience involved performing as a therapist for 30 minutes with a fourth-year student acting as the client. The therapy session represented the first session between the therapist and the client. The students acting as therapists received information about their patients' presenting problems several weeks in advance so that they could prepare adequately. The simulated therapy session was followed by a debriefing group session where the third-year students received feedback from their simulated clients and were guided through a reflection on the experience. During this reflective session, third-year students usually reported feeling anxious at the beginning of the simulated session and more secure later.

Statistical analysis

Means, standard deviations, skewness, and kurtosis were computed for the 29 TBS item scores. To test their factorial validity, two models from Emery et al. (2009) were specified and analyzed via CFAs using Structural Equation Modeling (SEM). The first was a four-factor model with 29 items loading onto four latent factors (Responsibility, Control, Inflexibility, and Distress), while the second was a three-factor model with 12 items loading onto three latent factors (Control, Inflexibility, and Distress). As Emery et al. (2009) did, a single-factor solution was also tested for comparison. As the variables were ordinal, the maximum likelihood with robust standard errors and a mean- and variance-adjusted (MLMV) estimation method was employed (DiStefano, 2002). Model fit was assessed using fit indexes and standard criteria ($\chi^2/df < 3$; CFI, TLI > 0.90; SRMR < 0.08; RMSEA < 0.08; Hair et al., 2016). As neither model showed acceptable fit, further examination involved inspecting results (e.g., item descriptive statistics, factor loadings, modification indices, content analysis, and factor correlations) and refining the four-factor model by discarding some items and reallocating others. This process involved an item descriptive statistics analysis and a content analysis of item allocation to factors conducted by the first, second, and last authors. Modification indices suggesting item reassignments were also inspected. An item was removed if its distribution indicated reduced discriminative power and informativeness (i.e., high skewness and kurtosis values). An item was reallocated to another factor if suggested by modification indices (i.e., > 3.84) and at least two out of three authors during content analysis. Factor loadings were examined to remove irrelevant items ($\lambda < .40$) and items showing cross-loadings across multiple factors (thus impeding specific factor measurement). Correlations between factors were also inspected to detect possible overlapping. Finally, to retain a similar number of items for each factor, the items with the lowest loadings were removed. This process led to a three-factor correlated model, and two alternative models were then tested: a hierarchical model, where the three factors loaded onto a second-order factor, and a bifactor model, where items loaded on a general factor in addition to specific factors.

McDonald's ω coefficients were calculated to assess the internal consistency of the CFA-derived dimension scores. Both direct and factorial scores were computed for each dimension, and their correlation was obtained. Descriptive statistics were then computed for these dimension scores. Criterion-related validity was examined by calculating Pearson's correlations between TBS scores, its dimensions, trait anxiety, anxiety sensitivity, optimism, and self-efficacy. Sensitivity analyses to study the TBS scores' ability to discriminate between groups involved computing their Pearson correlation with age and mean difference tests using *t*-tests for comparisons between two categories and ANOVA (Kruskal-Wallis if homoscedasticity could not be assumed) for comparisons between three or more, with appropriate post-hoc analyses (Tukey's HSD if homoscedasticity could be assumed, Games-Howell if not) followed accordingly. Effect sizes were also obtained, using Cohen's *d* for *t*-tests (Glass' δ if homoscedasticity could not be assumed) and η^2 for ANOVAs and Kruskal-Wallis' tests.

Materials and analysis code for this study are available by emailing the corresponding author. This study was approved by the Institutional Ethics Committee (approval reference: CEI21-02) but was not preregistered in a publicly accessible registry system. MPlus 7.2 (Muthén & Muthén, 2012) was used for SEM analyses and SPSS 27 for the rest.

Results

Item descriptive statistics

Table 2 presents descriptive statistics for the 29 TBS item scores, showcasing means predominantly between 2 and 4 within a response range of 1-6. Standard deviations ranged from .90 to 1.50, indicating similar variability across items. Most items exhibited normal skewness and kurtosis values, except for item 4, where a noteworthy 46.3% of responses were 1 (Strongly disagree). This item displayed moderate positive skewness, signifying data clustering below the mean, and high kurtosis, indicating a very peaked curve. These findings suggest a non-normal distribution, signaling reduced discriminative power and informativeness.

Table 2. Descriptive statistics of the scale items

Item	M	SD	Skewness	Kurtosis	CTIC
1. I must not make mistakes in therapy if I do then I've failed	2.70	1.22	.32	-.61	.45
2. There is no room for mistakes in therapy	2.75	1.27	.30	-.65	.42
3. If my clients do not progress it is my responsibility	3.15	1.07	-.11	-.52	.49
4. If I just stick to one therapeutic model it will solve the problem for me	1.77	.93	1.68	4.19	.23
5. Once I have decided on a treatment model I should stick to it	2.46	1.21	.67	-.07	.43
6. I am responsible if therapy is not successful	2.96	1.05	.02	-.51	.50
7. I must fully understand my client or I won't feel effective	3.60	1.22	-.28	-.49	.51
8. If I don't understand what happens in therapy I'm extremely uncomfortable	4.01	1.21	-.58	.14	.30
9. The course of therapy should be predictable	2.82	1.05	.23	-.30	.37
10. If I deviate from the clinical model then I've failed	2.24	1.03	.74	.51	.45
11. If I allow my clients to distress me I'm a failure	2.40	1.09	.60	-.08	.50
12. I don't need to fully understand what happens in therapy in order to help the client	3.07	1.24	.21	-.70	-.27
13. I must work at peak efficiency at all times	4.12	1.23	-.59	-.11	.48
14. I should be able to achieve results within a certain time frame	3.70	1.13	-.40	-.06	.36
15. Strong emotions will overwhelm and damage my client because they are fragile	2.39	1.15	.57	-.22	.52
16. I must always adhere strictly to a therapeutic model or I'm unprofessional	2.60	1.17	.51	-.25	.45
17. If I don't have all the information, I'm uncomfortable with therapy	3.49	1.07	-.27	-.17	.43
18. I must protect my client from reliving painful events	2.64	1.28	.61	-.28	.46
19. If I work hard enough therapy will always be successful	2.61	1.17	.49	-.24	.43
20. If I allow myself to feel what my client feels I'll be damaged	3.30	1.28	.07	-.56	.38
21. I must not allow my client to become too distressed in therapy	2.97	1.27	.28	-.69	.51
22. If I am affected by my client's story it will paralyse me and make me ineffective	3.32	1.18	.05	-.59	.35
23. If I have strong reactions to my clients, it means I'm abnormal	2.19	.96	.76	.97	.35
24. I shouldn't allow my clients to become distressed, they really want to feel better	2.56	1.18	.63	.06	.57
25. It is unprofessional to take an eclectic approach to therapy	2.51	1.17	.58	.16	.41
26. It is my role to find the solutions to my client's difficulties	3.23	1.46	.03	-1.00	.48
27. It is unprofessional to act spontaneously in therapy	2.38	1.04	.57	.24	.44
28. I should treat all clients the same way	2.28	1.43	1.05	.17	.24
29. I should be emotionally available to my client at all times	2.52	1.29	.60	-.37	.40

Note. CTIC = corrected total-item correlation. The items in the final solution were the following: Items 1, 2, 3, 6, and 7 for Responsibility-Control, items 5, 10, 16, and 25 for Inflexibility, and items 15, 18, 21, and 24 for Distress. The Spanish translation of all items is available in Appendix 1.

Confirmatory factor analyses

The CFAs comparing Emery et al.'s (2009) models, i.e., the four-factor model with 29 items loading on four correlated factors and the three-factor model with 12 items loading on three correlated factors, revealed poor fits for both models (see Table 3). Additionally, the one-factor solution showed an even worse fit compared to the previous models.

Table 3. Goodness of fit indices of the factorial solutions tested through CFA

Model	χ^2	df	p	χ^2/df	CFI	TLI	RMSEA	SRMR
Original 4-factor	1038.84	371	< .001	2.80	.71	.68	.06	.08
Original 3-factor	173.75	51	< .001	3.41	.88	.84	.07	.07
One-factor	1244.37	377	< .001	3.30	.62	.60	.07	.08
Depurated correlated	90.20	56	< .001	1.61	.98	.97	.03	.04
Depurated hierarchical	90.20	56	< .001	1.61	.98	.97	.03	.04
Depurated bifactor	66.54	46	.03	1.45	.99	.98	.03	.03

Note. The original 4-factor model had 29 items. The original 3-factor model had 12 items. The depurated model has 13 items. *df* = degrees of freedom. CFI = comparative fit index. TLI = Tucker-Lewis index. RMSEA = root mean square error of approximation. SRMR = standardized root mean squared residual.

Given the inadequate fit of these models, the four-factor model underwent refinement. This involved item content analysis and inspection of item descriptive statistics, modification indices, factor loadings, and factor correlations, as described in the Statistical analysis section. Following these adjustments, item 4 was removed because of its aforementioned skewed and peaked distribution (see Table 2). Items 1 and 2 were shifted from Distress to Control and Responsibility, respectively. Items 8, 12, 20, 22, 23, 28 were removed due to low factor loadings ($\lambda < .40$). The Responsibility and Control factors, highly correlated at all steps ($r = .89$-.91), were merged into a single factor labeled Responsibility-Control. Items 11, 19, 26, 27, and 29 were also removed as they showed similar loadings across multiple factors, impeding specific factor measurement. At this point, the Distress and Inflexibility factors each held four items, while Responsibility-Control retained nine. To maintain balance, the four items with the lowest loadings on Responsibility-Control (items 9, 13, 14, and 17) were removed. This three-factor correlated model exhibited excellent fit, as evidenced in Table 3. Distress was correlated with Inflexibility ($r = .50$) and Responsibility-Control ($r = .48$), and the latter two factors were also associated ($r = .58$). Two alternative models were formulated, both with 13 items and three factors. The first, a hierarchical model with the three factors loading on a second-order factor (TB), demonstrated the same fit as the correlated solution. The second, a bifactor model with all items loading on a general factor (TB) in addition to their specific factors, displayed superior fit and was consequently retained for subsequent analyses. Figure 1 illustrates the structure of this solution, showcasing standardized factor loadings.
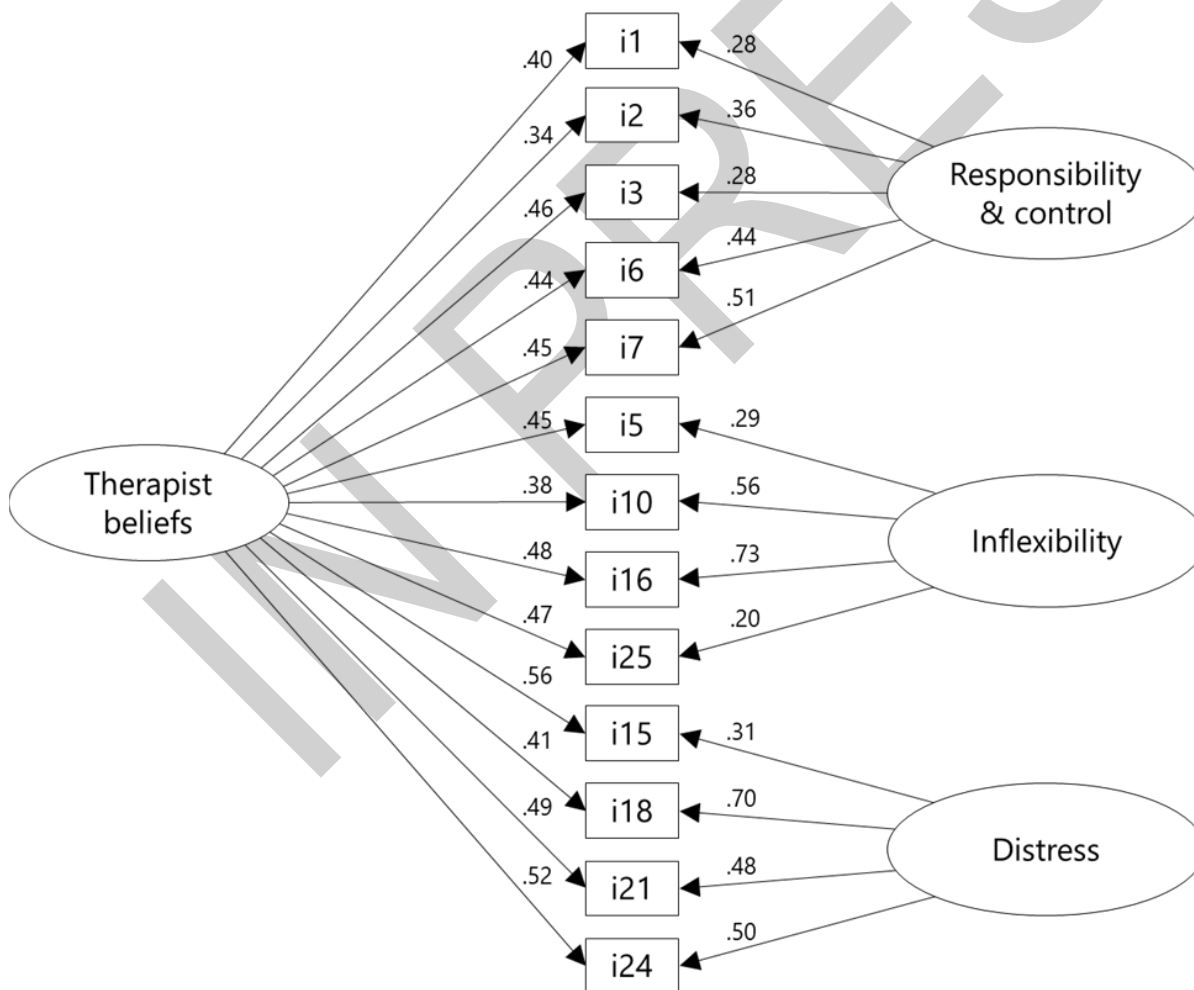

Figure 1. Bifactor final standardized solution for the Therapist Belief Scale (TBS). TB = therapist beliefs general factor.

10

Reliability and descriptive statistics of the dimensions

McDonald's ω coefficients indicated adequate internal consistency for all TBS and subscale scores (TBS, ω = .86; Distress, ω = .80; Inflexibility, ω = .74; Responsibility-Control, ω = .70).

Observed and factorial scores were computed, and their correlations were very strong in all cases ($r = .99$ for Responsibility-Control, $r = .97$ for Inflexibility, $r = .99$ for Distress, and $r = .998$ for general TB). Consequently, observed scores were utilized for subsequent analyses due to their equivalency to factorial scores and ease of computation for applied settings. The mean scores for each dimension were as follows: Responsibility-Control, $M = 15.16$ ($SD = 4.20$); Inflexibility, $M = 9.80$ ($SD = 3.31$); Distress, $M = 10.56$ ($SD = 3.92$); and TB, $M = 35.52$ ($SD = 8.70$).

Evidence of criterion-related validity

Table 4 displays the correlations between TB and psychological variables. Greater TB endorsement, both globally and on the Responsibility-Control and Distress dimensions, was significantly correlated with higher anxiety sensitivity and trait anxiety, and lower optimism and self-efficacy ($p < .001$ in all cases but one, see Table 4). These correlations were weak for Distress and moderate for Responsibility-Control and the overall TB score. Inflexibility exhibited a significant ($p < .05$) yet weak association with anxiety sensitivity (positive) and self-efficacy (negative), while no links emerged to trait anxiety or optimism ($p > .05$).

Table 4. Pearson correlation coefficients between therapist beliefs and other variables

Therapist beliefs	Anxiety sensitivity	Trait anxiety	Optimism	Self-efficacy
Total	.22***	.22***	-.19***	-.23***
Responsibility-Control	.24***	.25***	-.23***	-.26***
Inflexibility	.09*	.07	-.07	-.09*
Distress	.15***	.16***	-.12**	-.16***

Note. *** $p < .001$. ** $p < .01$. * $p < .05$

Evidence of sensitivity

Concerning the relations with sociodemographic, study, and practice-related variables (see Table 1), both the overall TB score and its three dimensions showed a negative and weak correlation with age. Gender differences were observed, with women displaying lower scores than men across the overall score and the Inflexibility and Distress dimensions, with small effect sizes. Participants' level of study also revealed significant differences: undergraduates reported higher scores on the global scale and its three dimensions compared to master's degree students and practicing psychologists (with a small effect size for Inflexibility, medium for Responsibility-Control, and large for Distress and the overall scale). Additionally, fourth-year undergraduate students (who had undertaken an internship) displayed significantly lower mean scores on Distress ($M = 9.36$, $t[223] = 5.86$, $p < .001$, $d = .88$; large effect size), Responsibility-Control ($M = 14.48$, $t[223] = 3.58$, $p < .001$, $d = .54$; medium effect size) and the overall TB score ($M = 33.66$, $t[137.10] = 5.75$, $p < .001$, $\delta = .94$; large effect size) compared to first-years (respectively, $M = 12.65$, $M = 16.67$, $M = 39.79$). There were no differences in Inflexibility between first- ($M = 10.46$) and fourth-year students ($M = 9.82$, $t[223] = 1.37$, $p = .17$, $d = .21$).

In terms of program modality, differences surfaced in Distress and the global scale: students in full attendance programs displayed lower scores compared to those in part attendance programs, with small effect sizes. Significant differences also emerged by work experience as psychologists: those who had worked reported lower scores, with small effect sizes for Inflexibility, medium for Responsibility-Control, large for Distress, and medium-large for the global score. Moreover, significantly lower scores in Distress and the global score were observed among students who had completed internships (large effect size) and those who received targeted TB

training (small effect size) compared to those who had not.

Regarding the TBS scores' sensitivity to detect changes before and after the therapist-client simulation in third-year students, Table 5 illustrates the TBS means and standard deviations at both timepoints, along with the results of *t*-tests. The mean scores for the overall TBS scale and its three dimensions were significantly lower after the simulation, with medium effect sizes for Responsibility-Control and the overall score and small for Inflexibility and Distress.

Table 5. Descriptive statistics of therapist beliefs and paired t-tests before and after the therapist-client simulation

Therapist beliefs	Pre-simulation		Post-simulation			Cohen's *d*
	M	*SD*	*M*	*SD*	*p*	
Total	32.70	7.28	29.16	8.32	< .001	.66
Responsibility-Control	14.51	3.76	12.53	4.00	< .001	.67
Inflexibility	8.77	2.90	8.16	3.01	.02	.25
Distress	9.42	2.99	8.47	3.22	< .001	.39

Discussion

This study focused on back-translating the TBS into Spanish and examining its scores' reliability and validity. In this process, item 4 was excluded due to its measurement inadequacy, and Emery et al. (2009) four- and three-factor solutions did not adequately represent TB in Spanish psychologists and psychology students. Through meticulous analysis and refinement, following theoretical and empirical criteria, a 13-item bifactor model emerged, with three specific factors (Distress, Inflexibility, and Responsibility-Control) and one general TB factor. This final model exhibited exceptional fit indices and internal consistency analyses underscored the TBS scores' reliability. Notably, correlations between direct and factorial scores across all factors suggested minimal information loss with the use of direct scores, bolstering the practicality of their application.

This study compellingly established criterion-related validity by revealing significant correlations between various TBS scores and psychological variables. Notably, the global TB, Responsibility-Control, and Distress scores exhibited positive associations with anxiety sensitivity and trait anxiety, aligning with previous research linking these psychological constructs to a low tolerance for uncertainty and need for control (Roscoe & Taylor, 2023). Additionally, Inflexibility showed a significant positive relationship with anxiety sensitivity, in line with previous findings (D'Souza Walsh et al., 2019; Simpson-Southward et al., 2018). The negative correlations observed between the global TB, Responsibility-Control, and Distress scores and optimism (Hoy et al., 2008), as well as self-efficacy (Lent et al., 2009; Emery et al., 2009), replicated previous findings and further supported the criterion-related validity of the TBS scores.

In analyzing the TBS scores' sensitivity across various variables, gender differences revealed women scoring lower than men on Inflexibility, Distress, and the global TB scale, consistent with previous studies highlighting greater uncertainty tolerance among female healthcare workers (Carney et al., 2004; Carney et al., 2007; McCulloch et al., 2005). Given that beliefs impact psychological practice (Altabef et al., 2017), this could signify that female therapists have a more positive attitude toward their patients (Artkoski & Saarnio, 2013). Age displayed negative correlations with the three TBS dimensions and the overall score, mirroring previous findings (Emery et al., 2009), possibly tied to study level and experience. Indeed, undergraduate students reported higher TBS scores than master's degree students and working psychologists, particularly on Distress and the global score, aligning with prior literature (Altabef et al., 2017; Green et al., 2017; Hill et al., 2007; Morrison & Lent, 2018). As expected, fourth-year undergraduates scored significantly lower on the global TB scale than first-years (Falanga et al., 2014). Moreover, specific training in TB was associated with lower Distress and global TB scores,

supporting Tyron's (2001) findings on reduced worry and increased confidence with formal training.

Regarding practical experience, participants who had worked as a psychologist scored lower in all TB dimensions and the global scale, while students who had completed internships scored lower in Distress and the global TB scale. Likewise, third-year students' global TBS scores and dimension scores significantly decreased after the therapist-client simulation. These findings align with past studies emphasizing the impact of practical and self-reflective experiences (Bennett-Levy & Finley-Jones, 2018; Chigwedere et al., 2021; Falanga et al., 2014; Haarhoff et al., 2015; Haarhoff & Kazantzis, 2007; Lent et al., 2006; Swift et al., 2018). Additionally, students in full attendance programs scored lower than those in part attendance programs, corroborating Abrams et al.'s (2017) study on counselor exam scores.

The sensitivity findings discussed support the Spanish TBS scores' ability to differentiate between groups, complementing the evidence for reliability, factorial, and criterion-related validity and bolstering the scale's credibility in assessing TB among psychology professionals and students.

This study's strengths lie in the back-translation method employed to adapt the TBS to Spanish and the advanced methods used to explore its factor structure (i.e., CFAs through SEM). Additionally, the comprehensive investigation of multiple validity aspects—reliability, factor validity, criterion-related validity, and sensitivity—contributes significantly to understanding the psychometric properties of the TBS scores. By refining the measurement of TB through this adapted instrument, this study has not only demonstrated its reliability and validity but also provided a tool that can be used to further explore how therapist beliefs interact with broader therapist effects. In fact, some variables considered in this study, such as clinical experience, gender, or continuing education, are central and potentially influential professional outcomes. Other variables —for instance, the therapeutic model or a therapist's specific personal style— are also key in the study of therapist effects (Casari et al., 2019; Tabullo et al., 2023) and in therapist education. For example, therapists in training might benefit from learning how their theoretical orientation and personal characteristics may influence their therapeutic style and how this may help them adjust their approach to better meet the needs of their patients.

However, several limitations should be acknowledged. First, one of the main limitations of this study was that a significant proportion of the sample consisted of undergraduate students, many of whom may not necessarily be interested in clinical psychology or in the practice of psychotherapy. Therefore, the results should be interpreted with caution. The second limitation involves the limited representation of psychology professionals and other categories in the sample (e.g., there was only one Ph.D. student), as well as a gender imbalance that, while consistent with the psychological field and prior TBS studies (Emery et al., 2009), may impact generalizability to men and other genders. A larger and more diverse sample would enable deeper analyses, including the cross-validation of the factor structure across psychology students and professionals. Additionally, although a diverse range of types of validity evidence was studied, future research should continue to gather information on the psychometric properties of the Spanish TBS scores, not only replicating the present results but also examining other sources of validity, such as the temporal stability of the scores or their convergent validity with other equivalent instruments. Finally, while the Spanish TBS scores demonstrated validity with psychology students and professionals, further research is needed to explore its applicability across other mental health professionals (e.g., psychiatrists, social workers, occupational therapists) and healthcare providers (e.g., physicians, nurses, physiotherapists) for a more comprehensive understanding of TB and their links to performance and therapy outcomes.

This study holds significant theoretical and practical implications. First, it underscores the need for more extensive research on TB within the Spanish context, a relatively unexplored area. Considering its potential ties to therapist well-being, professional performance, and therapy outcomes, future research should delve into causal relationships between these variables and their generalizability across different intervention settings and populations. Second, the information provided by the TBS could aid in identifying supervision needs among therapists, guiding effective interventions to address these needs. Incorporating regular assessments of TB into training and supervision could enhance therapists' self-awareness and ultimately improve treatment outcomes. Such assessments may help identify areas where targeted interventions—such as self-reflective practices,

debriefings, and focused training modules—can mitigate maladaptive beliefs and foster more adaptive therapeutic attitudes. This, in turn, could contribute to more effective psychotherapy by ensuring that therapists are not only technically proficient but also attuned to the personal and professional factors that influence their performance.

Third, assessing TB in students could provide a longitudinal perspective on skill acquisition, crucial for understanding the underlying processes of change, and guide the development of TB training while providing insights into its effectiveness. Different learning modalities should be investigated, as students in blended learning programs in particular seem to face unique challenges related to TB, suggesting the need for tailored study programs. Therefore, it is crucial for universities and training programs to integrate TB training into their curricula. The importance of experiential learning actions, as highlighted by the impact of training on TB (Bennett-Levy & Finley-Jones, 2018; Chigwedere et al., 2021; Haarhoff et al., 2015; Haarhoff & Kazantzis, 2007), underscores the need to incorporate practices such as internships and therapist-client simulations. These actions should place a strong emphasis on promoting self-reflection (e.g., through debriefings), enabling psychology students to approach clinical practice and enhance their empathy, confidence, competence, and meta-competence (Freeston et al., 2019; Gale & Schroder, 2014).

In conclusion, this study introduces a relevant tool for assessing TB within the Spanish context. Through CFA, the study establishes a well-defined scale consisting of three specific factors (Distress, Inflexibility, and Responsibility-Control) alongside a broader TB factor. The findings support the reliability and validity of the scale scores, considering different validity sources. Consequently, this research paves the way for diverse research avenues and practical settings, such as informing psychotherapy training and practice.

Conflict of Interest

The authors have no conflicts of interest to disclose.

References

Abrams, C. M., de Raet, A., Mcculloch, D., Elliott, G., & Meder, C. (2017). The impact of program modality on Counselor Preparation Comprehensive Exam subscale scores. *Journal of Counselor Preparation and Supervision*, *10*(1), Article 5. https://digitalcommons.sacredheart.edu/jcps/vol10/iss1/5

Alavi, M., Visentin, D. C., Thapa, D. K., Hunt, G. E., Watson, R., & Cleary, M. (2020). Exploratory factor analysis and principal component analysis in clinical studies: Which one should you use. *Journal of Advanced Nursing*, *76*(8), 1886-1889. https://doi.org/10.1111/jan.14377

Altabef, D., Meier, S., Reynolds, A., Delucia, J., & Friedling, L. (2017). Therapist response to a distressed client: Differences in active listening and changes in negative affect. *Counselling and Psychotherapy Research, 17*(3), 234–239. https://doi.org/10.1002/capr.12124

Artkoski, T., & Saarnio, P. (2013). Therapist's Gender and Gender Roles: Impact on Attitudes toward Clients in Substance Abuse Treatment. *Journal of Addiction*, *2013*, Article 591521. https://doi.org/10.1155/2013/591521

Bennett-Levy, J. (2019). Why therapists should walk the talk: The theoretical and empirical case for personal practice in therapist training and professional development. *Journal of Behavior Therapy and Experimental Psychiatry, 62*, 133–145. https://doi.org/10.1016/j.jbtep.2018.08.004

Bennett-Levy, J., & Finlay-Jones, A. (2018). The role of personal practice in therapist skill development: A model to guide therapists, educators, supervisors and researchers. *Cognitive Behaviour Therapy, 47*(3), 185–205. https://doi.org/10.1080/16506073.2018.1434678

Bennett-Levy, J., McManus, F., Westling, B. E., & Fennell, M. (2009). Acquiring and refining CBT skills and competencies: Which training methods are perceived to be most effective? *Behavioural and Cognitive Psychotherapy*, *37*(5), 571–583. https://doi.org/10.1017/S1352465809990270

Brosan, L., Reynolds, S., & Moore, R. G. (2008). Self-evaluation of cognitive therapy performance: Do therapists know how competent they are? *Behavioural and Cognitive Psychotherapy, 36*(5), 581–587. https://doi.org/10.1017/s1352465808004438

Cano-García, F. J., Sanduvete-Chaves, S., Chacon-Moscoso, S., Rodríguez-Franco, L., García-Martínez, J., Antuna-Beller, M. A., & Pérez-Gil, J. A. (2015). Factor structure of the Spanish version of the Life Orientation Test-Revised (LOT-R): Testing several models. *International Journal of Clinical and Health Psychology, 15*(2), 139–148. https://doi.org/10.1016/j.ijchp.2015.01.003

Carney, P. A., Elmore, J. G., Abraham, L. A., Gerrity, M. S., Hendrick, R. E., Taplin, S. H., Barlow, W. E., Cutter, G. R., Poplack, S. P., & D'Orsi, C. J. (2004). Radiologist uncertainty and the interpretation of screening. *Medical Decision Making*, *24*(3), 255–264. https://doi.org/10.1177/0272989X04265480

Carney, P. A., Yi, J. P., Abraham, L. A., Miglioretti, D. L., Aiello, E. J., Gerrity, M. S., Reisch, L., Berns, E. A., Sickles, E. A., & Elmore, J. G. (2007). Reactions to uncertainty and the accuracy of diagnostic mammography. *Journal of General Internal Medicine*, *22*, 234–241. https://doi.org/10.1007/s11606-006-0036-9

Casari, L. M., Ison, M. S., y Gómez, B. (2019). Estilo personal del terapeuta: comparación entre profesionales que trabajan en instituciones de adicciones y profesionales del área clínica general. *Revista Costarricense de Psicología, 38*(1), 57–73. https://doi.org/10.22544/rcps.v38i01.04

Castonguay, L. G., & Hill, C. E. (Eds.). (2017). *How and why are some therapists better than others? Understanding therapist effects.* American Psychological Association. https://doi.org/10.1037/0000034-000

Chigwedere, C., Bennett-Levy, J., Fitzmaurice, B., & Donohoe, G. (2021). Personal practice in counselling and CBT trainees: The self-perceived impact of personal therapy and self-practice/self-reflection on personal and professional development. *Cognitive Behaviour Therapy, 50*(5), 422–438. https://doi.org/10.1080/16506073.2020.1846608

Corbella, S., Balmaña, N., Fernández-Álvarez, H., Saúl, L. A., Botella, L., y García, F. (2009). Estilo personal

del terapeuta y teoría de la mente. *Revista Argentina de Clínica Psicológica, 18*(2), 125–133. https://doi.org/10.24205/03276716.2018.1082

D'Souza Walsh, K., Davies, L., Pluckwell, H., Huffinley, H., & Waller, G. (2019). Alliance, technique, both, or more? Clinicians' views on what works in cognitive-behavioral therapy for eating disorders. *International Journal of Eating Disorders, 52*(3), 278–282. https://doi.org/10.1002/eat.23033

DiStefano, C. (2002). The impact of categorization with confirmatory factor analysis. *Structural Equation Modeling: A Multidisciplinary Journal*, *9*(3), 327–346. https://doi.org/10.1207/S15328007SEM0903_2

Emery, S., Wade, T. D., & McLean, S. (2009). Associations among therapist beliefs, personal resources and burnout in clinical psychologists. *Behaviour Change, 26*(2), 83–96. https://doi.org/10.1375/bech.26.2.83

Falanga, R., De Caroli, M. E., & Sagone, E. (2014). Self-efficacy beliefs and representation of self and psychology a longitudinal study with psychology university students. *Procedia-Social and Behavioral Sciences*, *140*, 295–299. https://doi.org/10.1016/j.sbspro.2014.04.423

Fernández-Castro, J., Álvarez, M., Blasco, T., Doval, E., y Sanz, A. (1998). Validación de la Escala de Competencia Personal de Wallston: Implicaciones para el estudio del estrés [Validation of the Wallston's Personal Competence Scale and its implication for research into stress]. *Ansiedad y Estrés, 4*(1), 31–41. https://www.ansiedadyestres.es/sites/default/files/rev/ucm/1998/anyes1998a3.pdf

Ferrando, P. J., Chico, E., y Tous, J. M. (2002). Propiedades psicométricas del test de optimismo Life Orientation Test [Psychometric properties of the 'Life Orientation Test' (LOT): An empirical application of the Multidimensional Graded Response Model]. *Psicothema, 14*(3), 673–680. https://www.psicothema.com/pi?pii=782

Freeston, M., Thwaites, R., & Bennett-Levy, J. (2019). 'Courses for Horses': designing, adapting and implementing self-practice/self-reflection programmes. *The Cognitive Behaviour Therapist, 12*, Article e28. https://doi.org/10.1017/S1754470X19000138

Gale, C., & Schroder, S. (2014). Experiences of self-practice/self-reflection in cognitive behavioural therapy: A meta-synthesis of qualitative studies. *Psychology and Psychotherapy: Theory, Research and Practice, 87*(4), 373–392. https://doi.org/10.1111/papt.12026

Green, R. A., Conlon, E. G., & Morrissey, S. A. (2017). Task values and self‐efficacy beliefs of undergraduate psychology students. *Australian Journal of Psychology, 69*(2), 112–120. https://doi.org/10.1111/ajpy.12125

Guillén-Riquelme, A., y Buela-Casal, G. (2013). Versión Breve del STAI en Adolescentes y Universitarios Españoles [STAI Short Version in Adolescents and College Spanish]. *Terapia Psicológica, 31*(3), 293–299. http://dx.doi.org/10.4067/S0718-48082013000300004

Haarhoff, B., & Kazantzis, N. (2007). How to supervise the use of homework in cognitive behavior therapy: The role of trainee therapist beliefs. *Cognitive and Behavioral Practice, 14*(3), 325–332. https://doi.org/10.1016/j.cbpra.2006.08.004

Haarhoff, B., Thwaites, R., & Bennett‐Levy, J. (2015). Engagement with self-practice/self-reflection as a professional development activity: The role of therapist beliefs. *Australian Psychologist, 50*(5), 322–328. https://doi.org/10.1111/ap.12152

Hair, J. F., Hult, G. T. M., Ringle, C., & Sarstedt, M. (2016). *A primer on partial least squares structural equations modeling (PLS-SEM)*. Sage.

Heinonen, E., Lindfors, O., Laaksonen, M. A., & Knekt, P. (2012). Therapists' professional and personal characteristics as predictors of outcome in short- and long-term psychotherapy. *Journal of Affective Disorders, 138*, 301-312. https://doi.org/10.1016/j.jad.2012.01.023

Heinonen, E., & Nissen-Lie, H. A. (2019). 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

Hill, C. E., Sullivan, C., Knox, S., & Schlosser, L. Z. (2007). Becoming psychotherapists: Experiences of novice

trainees in a beginning graduate class. *Psychotherapy: Theory, Research, Practice, Training, 44*(4), 434–449. https://doi.org/10.1037/0033-3204.44.4.434

Hoy, A. W., Hoy, W. K., & Kurz, N. M. (2008). Teachers' academic optimism: The development and test of a new construct. *Teaching and Teacher Education, 24*(4), 821–835. https://doi.org/10.1016/j.tate.2007.08.004

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. https://doi.org/10.1016/j.cpr.2018.08.004

Lent, R. W., Cinamon, R. G., Bryan, N. A., Jezzi, M. M., Martin, H. M., & Lim, R. (2009). Perceived sources of change in trainees' self-efficacy beliefs. *Psychotherapy: Theory, Research, Practice, Training, 46*(3), 317–327. https://doi.org/10.1037/a0017029

McCulloch, P., Kaul, A., Wagstaff, G. F., & Wheatcroft, J. (2005). Tolerance of uncertainty, extroversion, neuroticism and attitudes to randomized controlled trials among surgeons and physicians. *British Journal of Surgery, 92*(10), 1293–1297. https://doi.org/10.1002/bjs.4930

McLean, S., Wade, T. D., & Encel, J. S. (2003). The contribution of therapist beliefs to psychological distress in therapists: An investigation of vicarious traumatization, burnout and symptoms of avoidance and intrusion. *Behavioural and Cognitive Psychotherapy, 31*(4), 417–428. https://doi.org/10.1017/S135246580300403X

Morrison, M. A., & Lent, R. W. (2018). The working alliance, beliefs about the supervisor, and counseling self-efficacy: Applying the relational efficacy model to counselor supervision. *Journal of Counseling Psychology, 65*(4), 512–522. https://doi.org/10.1037/cou0000267

Muthén, L. K., & Muthén, B. O. (2012). *Mplus User's Guide: Statistical Analysis with Latent Variables (7th ed.)*. Muthén & Muthén.

Rameswari, T., Hayes, B., & Perera-Delcourt, R. (2021). Measuring therapist cognitions contributing to therapist drift: a qualitative study. *The Cognitive Behaviour Therapist, 14*, Article e7. https://doi.org/10.1017/s1754470x21000039

Roscoe, J., & Taylor, J. (2023). Maladaptive therapist schemas in CBT practice, training and supervision: A scoping review. *Clinical Psychology & Psychotherapy, 30*(3), 510–527. https://doi.org/10.1002/cpp.2802

Sandín, B., Valiente, R. M., Chorot, P., y Santed, M. A. (2007). ASI-3: Nueva escala para la evaluación de la sensibilidad a la ansiedad. *Revista de Psicopatología y Psicología Clínica, 12*(2), 91–104. https://doi.org/10.5944/rppc.vol.12.num.2.2007.4036

Scheier, M. F., Carver, C. S., & Bridges, M. W. (1994). Distinguishing optimism from neuroticism (and trait anxiety, self-mastery, and self-esteem): A reevaluation of the Life Orientation Test. *Journal of Personality and Social Psychology, 67*(6), 1063–1078. https://doi.org/10.1037/0022-3514.67.6.1063

Simpson-Southward, C., Waller, G., & Hardy, G. (2018). Supervisor practice when guiding therapists working with depression: The impact of supervisor and patient characteristics. *The Cognitive Behaviour Therapist, 11*, Article e9. https://doi.org/10.1017/S1754470X18000089

Swift, J., Callahan, J.L., Cooper, M., & Parkin, S. R. (2018). The impact of accommodating client preference in psychotherapy: A meta-analysis. *Journal of Clinical Psychology, 74*(11), 1924–1937. https://doi.org/10.1002/jclp.22680

Tabullo, A. J., Casari, L. M., & Putrino, N. I. (2023). Associations between the personal style of the therapist, empathy and theory of mind among Argentine psychotherapists. *Revista de Psicopatología y Psicología Clínica, 28*(1), 51–65. https://doi.org/10.5944/rppc.33825

Taylor, S., Zvolensky, M. J., Cox, B. J., Deacon, B., Heimberg, R. G., Ledley, D. R., Abramowith, J. S., Holaway, R. M., Sandín, B., Stewart, S. H., Coles, M., Eng, W., Daly, E. S., Arrindell, W. A., Bouvard, M., & Jurado Cardenas, S. (2007). Robust dimensions of anxiety sensitivity: Development and initial validation of the Anxiety Sensitivity Index-3 (ASI-3). *Psychological Assessment, 19*(2), 176–188. https://doi.org/10.1037/1040-3590.19.2.176

17

Thériault, A., Gazzola, N., & Richardson, B. (2009). Feelings of incompetence in novice therapists: Consequences, coping, and correctives. *Canadian Journal of Counselling and Psychotherapy, 43*(2), 105–119. https://cjc-rcc.ucalgary.ca/article/view/58957

Tryon, G. S. (2001). School psychology students' beliefs about their preparation and concern with ethical issues. *Ethics & Behavior*, *11*(4), 375–394. https://doi.org/10.1207/S15327019EB1104_02

Van Knippenberg, F. C. E., Duivenvoorden, H. J., Bonke, B., & Passchier, J. (1990). Shortening the State-Trait Anxiety Inventory. *Journal of Clinical Epidemiology, 43*(9), 995–1000. https://doi.org/10.1016/0895-4356(90)90083-2

Waller, G., & Turner, H. (2016). Therapist drift redux: Why well-meaning clinicians fail to deliver evidence-based therapy, and how to get back on track. *Behaviour Research and Therapy, 77*, 129–137. https://doi.org/10.1016/j.brat.2015.12.005

Wallston, K. A. (1992). Hocus-Pocus, the focus isn't strictly on locus: Rotter's social learning theory modified for health. *Cognitive Therapy and Research, 16*(2), 183–199. https://doi.org/10.1007/BF01173488

Wampold, B. E., & Imel, Z. E. (2015). *The great psychotherapy debate: The evidence for what makes psychotherapy work* (2nd ed.). Routledge.

Appendix 1
Spanish version of the Therapist Belief Scale (TBS)

Instrucciones: Lee atentamente cada afirmación e indica tu grado de acuerdo con ella marcando la respuesta que corresponda, teniendo en cuenta las siguientes opciones:

1. Totalmente en desacuerdo	4. Algo de acuerdo
2. Muy en desacuerdo	5. Muy de acuerdo
3. Algo en desacuerdo	6. Totalmente de acuerdo

1. No debo de cometer errores en la terapia, y si los cometo, he fracasado. *
2. En terapia no se pueden cometer errores. *
3. Si mis clientes no avanzan, es mi responsabilidad. *
4. Me basta con adherirme a un único modelo terapéutico.
5. Una vez que me decido por un modelo de tratamiento, debería ceñirme a él. *
6. Si la terapia no tiene éxito, soy el/la responsable. *
7. Debo comprender plenamente a mi cliente o no me sentiré eficaz. *
8. Si no entiendo lo que ocurre en terapia, me siento extremadamente incómodo/a.
9. El curso de la terapia debería ser predecible.
10. Si me desvío del modelo clínico, he fracasado. *
11. Si permito que mis clientes me angustien, soy un fracaso.
12. No necesito entender completamente lo que sucede en la terapia para ayudar a mi cliente.
13. Tengo que trabajar con la máxima eficiencia en todo momento.
14. Debería ser capaz de obtener resultados en un determinado período de tiempo.
15. Las emociones intensas abrumarán y perjudicarán a mis clientes, ya que son frágiles. *
16. Siempre he de adherirme de forma estricta al modelo terapéutico o sería poco profesional. *
17. Si no tengo toda la información, me siento incómoda/o con la terapia.
18. He de proteger a mi cliente para que no reviva eventos dolorosos. *
19. Si me esfuerzo lo suficiente, la terapia funcionará siempre.
20. Si me permito sentir lo que mi cliente siente, me afectará.
21. No debo permitir que mi cliente se angustie demasiado en terapia. *
22. Si la historia personal de mi cliente me afecta, eso me paralizará y me hará ineficaz.
23. Si tengo reacciones intensas hacia mis clientes, implica que no soy normal.
24. No debería permitir que mis clientes se angustien, lo que buscan es sentirse mejor. *
25. No es profesional tener un enfoque ecléctico en la terapia. *
26. Mi rol es encontrar las soluciones para las dificultades de mi cliente.
27. Actuar de forma espontánea en terapia es poco profesional.
28. He de tratar a todos los clientes del mismo modo.
29. Debería de estar emocionalmente disponible para mi cliente en cualquier momento.

Note. The items in the final factorial solution are indicated with an asterisk (see Figure 1). The factors are structured as follows: *Responsibility and control*, comprising items 1, 2, 3, 6, and 7; *Inflexibility*, comprising items 5, 10, 16, and 25; and *Distress*, comprising items 15, 18, 21, and 24.