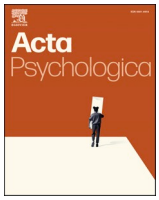
Acta Psychologica 263 (2026) 106272



Contents lists available at ScienceDirect

Acta Psychologica

journal homepage: www.elsevier.com/locate/actpsy





Development and validation of a Portuguese psychologist's empathy scale: A pilot study

Márcia Cruz [a], Ana Grilo [b], Susana Custódio [c], Artemisa R. Dores [d,*], Carina Silva [e], Paulo Alves [f]

[a] Escola Superior de Enfermagem do Porto, CINTESIS@RISE, Porto, Portugal
[b] H&TRC-Health & Technology Research Center, ESTeSL- Escola Superior de Tecnologia da Saúde, Instituto Politécnico de Lisboa, CICPsi - Research Center for Psychological Science, Faculdade de Psicologia, Universidade de Lisboa, Lisboa, Portugal
[c] ciTechCare - Centro de Inovação em Tecnologias e Cuidados da Saúde, Escola Superior de Saúde, Instituto Politécnico de Leiria, Leiria, Portugal
[d] CIR, Escola Superior de Saúde, Instituto Politécnico do Porto; Laboratory of Neuropsychophysiology, Faculty of Psychology and Education Sciences, University of Porto, Porto, Portugal
[e] ESTeSL- Escola Superior de Tecnologia da Saúde, Instituto Politécnico de Lisboa; Centro de Estatística e Aplicações, Faculdade de Ciências da Universidade de Lisboa, Portugal
[f] Insight - Piaget Research Center – ISEIT / Viseu, Portugal; CIEP – Centro de Investigação em Educação e Psicologia – Universidade de Évora, Portugal

ARTICLE INFO

Keywords:
Empathy
Psychologists
Scale development
Factor analysis
Psychometric properties

ABSTRACT

The systematic assessment of empathy is essential for identifying psychologists' competencies and developmental needs, ultimately contributing to more effective clinical practice. Given the multidimensional nature of empathy, comprehensive evaluation requires consideration of its cognitive, affective, and behavioural dimensions. However, no existing instrument is specifically validated to assess psychologists' self-perceived empathy in clinical contexts. To address this limitation, the present pilot study aimed to develop and undertake a preliminarily validation of the Portuguese Psychologists' Empathy Scale (PPES), a self-report instrument designed to assess empathy among psychologists working in Portuguese-speaking contexts across these three components. The study proceeded in two phases: item construction and initial psychometric testing. A total of 106 Portuguese psychologists completed a 57-item Likert-type instrument. Bartlett's test of sphericity was significant ($\chi^2(45) = 811.56$, $p < .001$), and the KMO index (0.73) indicated moderate adequacy for factor analysis. Exploratory factor analysis revealed three core factors explaining 48% of the total variance, with strong internal consistency. A subsequent confirmatory factor analysis supported model refinement, leading to the removal of items with low factor loadings. The final PPES comprises 43 items distributed across cognitive (16), affective (16), and behavioural (11) dimensions. The findings provide encouraging preliminary evidence for the PPES, demonstrating robust internal consistency and acceptable initial factorial validity. Nevertheless, further studies using independent samples and external validation procedures are required to fully establish its psychometric soundness and clinical applicability with Portuguese-speaking psychologists.

1. Introduction

Psychological interventions can be challenging, and according to the literature, therapist characteristics may influence intervention outcomes (Barkham et al., 2017). Moreover, various factors may lead healthcare professionals, including psychologists, to experience emotional overload, potentially reducing their capacity to manage emotions effectively. Nevertheless, the establishment of a strong therapeutic relationship, particularly through empathic understanding, appears to be one of the most powerful agents of change, enabling the therapist to comprehend the client's emotional experiences and communicate that understanding effectively (Grosseman & Dohms, 2021).

The literature increasingly highlights the links between empathy, personal values, and prosocial tendencies, offering a broader conceptual context for understanding empathy in the helping professions. Empathy has been shown to predict prosocial orientation and altruistic behaviour, with studies demonstrating its association with cooperative, helping, and socially responsible actions (Ishtiyaq et al., 2024; Peng et al., 2024;

* Corresponding author at: CIR, Escola Superior de Saúde, Instituto Politécnico do Porto (E2S – P.Porto), 4200-072, Porto, Portugal; Laboratory of Neuropsychophysiology, Faculty of Psychology and Education Sciences, University of Porto, 4200-135, Porto, Portugal.
E-mail address: artemisa@ess.ipp.pt (A.R. Dores).

https://doi.org/10.1016/j.actpsy.2026.106272
Received 8 October 2025; Received in revised form 25 November 2025; Accepted 15 January 2026
Available online 25 January 2026
0001-6918/© 2026 The Authors. Published by Elsevier B.V. This is an open access article under the CC BY-NC license (http://creativecommons.org/licenses/by-nc/4.0/).

Sharghi, 2023). This body of work suggests that empathy cannot be viewed in isolation from the ethical and motivational systems that guide individuals' responses to others, thereby reinforcing its relevance as a multidimensional construct with interpersonal and moral implications in professional contexts.

From this perspective, professional empathy involves more than understanding or resonating with clients' experiences - it requires the intentional regulation of affect, ethical sensitivity, and behavioural responsiveness that translate empathic understanding into therapeutic practice. Recent studies indicate that prosocial orientations, value-driven motivations, and responsiveness to others' needs play an important role in how empathy is enacted in practice, particularly in relational professions such as psychology (Gungordu et al., 2025; Man & Jing, 2025).

Psychologists exhibiting higher levels of empathy are more likely to build stronger and more positive relationships with their clients, fostering an atmosphere of acceptance that encourages emotional expression, understanding, and personal growth, while establishing the trust and sense of security necessary for successful interventions (Bellosta-Batalla et al., 2019; Elliott et al., 2018; Muñoz, 2020; Norcross & Lambert, 2018). The therapist's empathic responsiveness, when combined with mutual relational empathy, seems to enhance the intervention process (Riker, 2020) and contribute to positive therapeutic outcomes (Barrera & de La Motte de Broöns de Vauvert, 2024).

In addition to its intrinsic value in therapeutic practice, empathy also plays a crucial role in fostering inclusivity, cultural sensitivity, and social justice within the field of psychology. By acknowledging and validating the diverse experiences and perspectives of clients, psychologists can create a more inclusive and equitable therapeutic environment that promotes recovery and empowerment. Therefore, empathy has been increasingly recognised as a core competency in psychological practice across diverse cultural and professional contexts (Decety & Cowell, 2015; Gerdes & Segal, 2009).

Despite the widespread recognition of its importance, the definition of empathy remains a subject of debate, particularly regarding its dimensions. The classical definition of empathy proposed by Rogers (1957) describes it as “to sense the client's private world as if it were your own, but without ever losing the ‘as if’ quality — this is empathy” (Rogers, 1957, p. 99). It can be defined as a psychological ability that allows us to understand the inner state of another individual (Riker, 2020). It is not a static trait but a dynamic skill set that underpins effective, ethical, and context-sensitive professional practice (Cai et al., 2026; Habib et al., 2025). Empathy is not only a foundational element of an effective therapeutic relationship but also a deliberate and powerful clinical intervention. For psychologists, it functions as a catalyst for trust, reinforcing the client's sense of self, and facilitates deeper psychological exploration and growth. Accordingly, empathy is regarded as an essential professional competence in effective clinical practice (Elliot & al., 2018; Elliott et al., 2019).

While the cognitive and affective dimensions are the most widely recognised, both centring on internal processes typically assessed through self-report methodologies, a third component, the behavioural dimension, has gained increasing conceptual and empirical support (Clark et al., 2018; Davis, 2018; Decety & Jackson, 2004; Gerdes & Segal, 2009). Cognitive empathy involves the ability to understand and interpret another person's thoughts and perspectives - commonly conceptualised as perspective-taking (Hogan, 1969). It reflects the tendency to understand others' internal states or the state of understanding another person's internal state (Clark et al., 2018). Affective, in turn, entails the capacity to share and resonate with the emotions experienced by others (Hoffman, 1984; Mehrabian & Epstein, 1972). Clark et al. (2018) define affective empathy as the tendency to experience emotional states that align with those of another person - that is, the experience of an affective state congruent that is congruent with another individual's emotional condition. Behavioural empathy extends beyond the mere understanding and sharing of emotions. It encompasses the tendency to engage in affective and/or cognitive empathy, including behavioural mirroring and/or empathic communication behaviours (Clark et al., 2018). This component drives psychologists to actively address client distress and foster well-being, reflecting a compassionate response to another person's experience (Decety & Jackson, 2004). It represents the externalisation of concern for the other person's well-being through both verbal and non-verbal expressions that legitimise and validate the other person's feelings. Behavioural empathy therefore entails not only understanding the client's perspective and emotional state but also regulating one's own affective responses and communicating this understanding in a therapeutically effective manner (Gerdes & Segal, 2009). The absence of this behavioural component may limit the effectiveness of empathy, reducing it to a mere intellectual or affective exercise, without meaningful therapeutic impact (Clark et al., 2018).

The three-dimensional structure of empathy is supported by both classical and contemporary authors (Barrett-Lennard, 1981; Davis, 1994; Decety & Jackson, 2004; Decety & Yoder, 2017; Elliott et al., 2018; Gerdes et al., 2010), who argue that the behavioural expression of empathy constitutes the observable dimension through which empathy becomes therapeutically effective. Nevertheless, significant gaps remain in the assessment of empathy, particularly due to ongoing challenges in defining the construct and determining which dimensions should be included, as well as how they should be measured (Clarke et al., 2015; Simões et al., 2023).

Empathy has been assessed using various methodologies, including self-report and informant-report instruments, observational techniques and rating scales, experimental methods, and psychometric tests (Grainger et al., 2023). Among the psychometric tests, the Interpersonal Reactivity Index (IRI; Davis, 1980, 1983) and the Jefferson Scale of Empathy (JSE; Hojat et al., 2001 ; Magalhães et al., 2012) are two of the most widely used, but many others are available (e.g., Questionnaire of Cognitive and Affective Empathy - QCAE; Reniers et al., 2011, Portuguese version by Queirós et al., 2018). For example, the IRI is a self-report instrument consisting of 28 items, and four subscales - perspective taking, empathic concern, personal distress, and fantasy - designed to assess cognitive and affective dimensions of empathy (Davis, 1980, 1983). The cognitive dimension is captured through perspective-taking, whereas the affective dimension is assessed through the remaining subscales (Davis, 1980). The IRI has been used to measure empathy in a variety of populations, including students of psychology and other fields (Davis, 1983; Konstam et al., 2003; Limpo et al., 2010), medical students (Lee et al., 2023), and the general adult population (Coutinho et al., 2015; Maddaluno et al., 2022; Manarte & Andrade, 2018). The JSE is a tool designed to assess empathy, specifically in physicians and other healthcare professionals involved in patient care (Hojat et al., 2001; Magalhães et al., 2012). Despite being originally developed for use in medical settings, the JSE has since been adapted for medical students and the general population. The instrument consists of 20-items (Hojat et al., 2001; Hojat et al., 2002) organised into three factors: perspective taking, compassionate care, and walking in a patient's shoes (Hojat et al., 2018). The JSE emphasises clinical empathy, which involves cognitive understanding and communication of patients' experiences, rather than purely affective empathy (Voultsos et al., 2024).

Across these measures, empathy is conceptualized primarily in terms of its cognitive and affective dimensions. However, in professional contexts, particularly in psychology, empathy extends beyond internal understanding or emotional resonance. It includes an intentional behavioural component that translates empathic insight and internal processes into therapeutic communication, action, and professional conduct. This integrated perspective justifies the adoption of a three-dimensional model capable of representing the complex, applied, and relational nature of empathy as enacted by psychologists in real-world settings.

Furthermore, these instruments were developed for administration to the general population (Brett et al., 2023; Konrath et al., 2018)

healthcare professionals or students (Bohler et al., 2021; Wilczek-Rużyczka & Kupczyk, 2021) and are not specifically designed to capture the essence of empathy in the clinical context of psychologists, which limits their relevance and sensitivity to the unique demands of psychological practice (Gerdes et al., 2011; Murphy & Lilienfeld, 2019).

This specificity is of paramount importance, as the nature of empathy in psychological settings involves unique interactions and demands that are not fully addressed by the general population or student measures, reflecting a clear need for a content-specific and context-relevant instrument to accurately assess empathy in psychological settings (Simões et al., 2023). Additionally, most validated empathy measures are available only in English or a limited number of other languages, which constrains cross-cultural comparability and may introduce construct and linguistic bias when instruments are used without proper adaptation (Benlidayi, 2024; Falk & O'Hara, 2021).

Therefore, despite the centrality of empathy in clinical work, a notable lack of validated, profession-specific instruments persists for assessing empathy as it is enacted and perceived by psychologists (Decker et al., 2014). This gap has been highlighted in a number of conceptual reviews, which have called for a more nuanced, context-specific understanding and measurement of empathy tailored to clinical psychology (Elliott et al., 2018; Murphy & Lilienfeld, 2019), rather than continued dependence on generic instruments (Cabedo-Peris et al., 2021; Cuff et al., 2016). The development of an empathy assessment instrument specifically designed for psychologists, grounded in a three-dimensional approach to empathy and consistent with contemporary evidence on professional competence and therapeutic effectiveness, addresses the previously mentioned gap, fostering self-reflection on empathic skills and offering pathways for enhancing therapeutic interventions. Consequently, it may contribute to improving the quality of psychological care and promoting more effective client-therapist interactions.

Therefore, this pilot study aims to develop and undertake a preliminary validation of a Portuguese self-report scale that assesses psychologists' self-perception of empathy across its three core dimensions in clinical practice.

2. Materials and methods

The development of the Portuguese Psychologists' Empathy Scale (PPES) is part of a multicentre project involving psychologists from Portuguese-speaking countries (Portugal, Brazil, Angola, and Mozambique), that aims to adapt and validate empathy assessment instruments across Lusophone contexts. Portuguese is an official language in countries spanning four continents and is currently spoken by more than 260 million people worldwide, making it one of the most widely spoken languages globally (Camões Institute, 2023). Establishing an empathy measure specifically for Portuguese-speaking psychologists is therefore critical, not only to enhance measurement equivalence and facilitate cross-cultural research collaborations (Beaton et al., 2000), but also to strengthen the quality of clinical training, supervision, and ongoing professional development within Portuguese-speaking countries.

The development and validation of the new instrument were conducted in two main stages: scale development and preliminary psychometric validation. The pilot study protocol was approved by the Ethics Committee of the Piaget Institute (CEIP) - RECI.2021.03.

2.1. Development of the Portuguese psychologists` empathy scale (PPES)

2.1.1. Item generation

The development of the PPES item pool was grounded in a systematic, theory-driven process to ensure conceptual and empirical robustness. Item generation was informed by a comprehensive review of the existing literature on empathy, with particular attention to established quantitative measures and theoretical conceptualisations. The process adopted a three-dimensional model of empathy (i.e., cognitive [DC], affective [DA], and behavioural [DB]), reflecting empathy's recognition as a multidimensional construct and its role as a core professional competency in psychological practice. Two primary sources guided item development. First, items were adapted and reformulated from validated empathy scales, such as the Interpersonal Reactivity Index (IRI; Davis, 1980) and the Empathy Inventory–IE (Falcone et al., 2008), ensuring continuity with well-established constructs and allowing for comparability across measures. Second, new items were generated through systematic analysis of theoretical frameworks addressing the processes and manifestations of empathy in professional contexts. This dual-source approach aimed to build a comprehensive pool that captured both established conceptual dimensions and emerging perspectives on empathic responding.

A noted limitation of existing empathy scales is their predominant emphasis on cognitive and affective dimensions, often neglecting behavioural expressions of empathy. To address this gap, the research team implemented a structured brainstorming procedure to elicit and generate items representing behavioural manifestations of empathy in a deliberate and systematic way. Three pairs of authors were formed within the research team, with each pair responsible for independently generating items related to one of the three empathy dimensions. This division enabled a focused, in-depth exploration of each domain while ensuring conceptual consistency across dimensions. Each pair drafted an extensive list of potential items aligned with theoretical definitions and practical expressions of their assigned component. These preliminary items were subsequently reviewed, discussed, and refined in joint team meetings through an iterative process aimed at enhancing conceptual coherence, linguistic precision, and face validity.

For the behavioural dimension, item creation was guided by the understanding that empathy can be enacted through both non-verbal behaviours (e.g., touch, eye contact) and verbal behaviours (e.g., expressing understanding, adjusting behaviour, engaging in active listening). In the affective dimension, particular attention was given to capturing the psychologist's emotional responses (e.g., anger, joy). The cognitive dimension focused on processes such as information processing, comprehension, and perspective-taking (e.g., understanding, thinking, comprehending).

2.1.2. Initial item pool and expert review

The initial item pool comprised 55 items: 26 affective (DA), 17 cognitive (DC), and 12 behavioural (DB). To minimise potential response biases, items were balanced between positively and negatively worded statements. The final questionnaire was subsequently randomised to avoid sequence or pattern effects that could influence participants' responses. This ensured that items representing different dimensions appeared in an unpredictable sequence, thereby enhancing the instrument's methodological rigour and reducing the likelihood of dimension-based response tendencies.

This preliminary version was then reviewed by 12 experienced psychologists, who provided feedback on the clarity and comprehensiveness of each item. Participants also had the opportunity to provide additional suggestions at the end of the scale. Therefore, seven participants proposed minor adjustments to the Portuguese wording. These suggestions were reviewed by the panel of authors, who, by consensus, decided to include them in most cases. Additionally, five participants found the item “I feel disgusted by some of the clients' accounts” unclear due to the use of the term “disgusted.” They asked whether this feeling of repulsion was meant to be more affective or cognitive. Consequently, this item was replaced with the following two items: DA_27r: I feel embarrassed by some of the clients' accounts. Moreover, DC_6r: I try to distance myself when the client talks about their problems or concerns, avoiding thinking from their point of view. This decision was also made by consensus among the authors. Therefore, based on expert feedback, the final version of the PPES comprised 57 items: 27 DA, 18 DC and 12 dB (Appendix A).

2.1.3. Scale format and scoring

The final 57-item scale included both positively and negatively worded statements. Items were rated on a 5-point Likert scale (1 = Strongly Disagree, 5 = Strongly Agree). Some items were reverse-scored, as they reflected behaviours, emotions, or cognitions contrary to the expression of empathy. Other items were constructed using negative phrasing. Higher total scores on the scale indicated greater perceived empathic competence.

2.2. Participants

Data collection took place over a period of six months, from October 2023 to March 2024. The primary inclusion criterion was current professional practice as a psychologist. The pilot study employed a convenience sample comprising 106 psychologists. Of these, 82% were female, with a mean age of 41.8 ± 9.7 (range: 25–66) years, and an average of 15.3 ± 9.5 years in professional practice, 43% were married, 60% held a master's degree, 54% worked professionally in the northern region, as employees (59%) and in clinical settings (65%). The majority of these psychologists (88%) reported providing individual care and intervention in their professional practice (Table 1).

2.3. Statistical analysis

2.3.1. Data screening and pre-analysis

Descriptive statistics were used to examine response variability and missing data. To identify and remove those items with the most missing data, an individual descriptive item analysis was performed. A missing-value analysis was also carried out to verify if it was necessary to remove questionnaires from the dataset. Response frequencies for each dimension were also calculated by averaging negative, neutral and positive responses. Negatively worded items were reverse-scored for further analysis. Tests of sphericity and sample adequacy (Kaiser-Meyer-Olkin - KMO (> 0.6) and Bartlett's test) were applied to assess the suitability of conducting Exploratory Factor Analysis (EFA) and Confirmatory Factor Analysis (CFA) (Hair et al., 2019).

Table 1
Demographics of participants (n = 106); M = Mean; SD = standard deviation.

Variable	%	*M*	±*SD*
Age (range: 25–66)	–	41.8	±9.7
Sex			
Female	82	–	–
Male	18	–	–
Marital status			
Married or in a civil partnership	65	–	–
Single or Separated/Divorced	35	–	–
Academic qualifications			
Doctorate	21	–	–
Masters	60	–	–
Bachelor	19	–	–
Working years		15.3	±9.5
1–5 years	24	–	–
6–10 years	9	–	–
11–20 years	39	–	–
21–40 years	28	–	–
Country region			
Islands	5	–	–
North	53	–	–
Centre	31	–	–
South	11		
Labour relationship			
As employees	52	–	–
Self-employment	29	–	–
Both	19	–	–
Intervention settings			
Clinical settings	64	–	–
Educational settings	33	–	–
Organisational settings	3	–	–
Assistance and intervention			
Individual	87	–	–
Group	13	–	–

2.3.2. Reliability

The reliability analysis (internal consistency) was conducted using Cronbach's alpha (α), which indicates the extent to which the questionnaire items can be treated as measuring a single latent dimension (Cronbach, 1951). Values above 0.7 are considered adequate for a questionnaire (Bland & Altman, 1997), although some authors consider 0.6 to be acceptable. For the entire questionnaire, Cronbach's alpha should be at least 0.9 (Bland & Altman, 1997). However, the validity of this measure has been questioned, and several authors have suggested alternative measures (McNeish, 2018). In this study, we also calculated the average inter-item correlation (AIIC), which is independent of the number of items in a dimension and the sample size. This measure evaluates how well the items within a dimension correlate with each other, that is, whether there is evidence that the items measure the same construct. A rule of thumb is that AIIC should be above 0.5 for good internal consistency, while values between 0.15 and 0.5 indicate moderate consistency (Clark & Watson, 1995). The composite reliability index (CR) is another measure of internal consistency, similar to Cronbach's alpha, but considered a more robust and appropriate measure in Structural Equation Modelling (SEM), especially when factor loadings vary significantly. Hair et al. (2019) suggested threshold of 0.70 as the minimum acceptable value for this index.

2.3.3. Theoretical validity

Theoretical validity is divided into three types: *factorial validity*, *convergent validity*, and *discriminant validity*. Factorial validity assesses whether the items reflect the dimension they are intended to measure and is generally evaluated using standardised factor loadings (λ_{ij}) (Child, 2006). A construct is said to have factorial validity if all items have $\lambda_{ij} \geq 0.5$, though $\lambda_{ij} \geq 0.2$ is considered acceptable (Child, 2006). Convergent validity is confirmed if the items within a dimension exhibit a high positive correlation (Fornell & Larcker, 1981). The Average Variance Extracted (AVE) was used to assess this validity, where values above 0.5 are considered appropriate. Discriminant validity ensures that the items measuring a specific dimension are not correlated with other constructs. The Fornell-Larcker criterion (Fornell & Larcker, 1981) specifies that the square root of the AVE of a dimension must be greater than its correlation with any other dimension in the model.

The dimensionality and construct validity were calculated via exploratory factor analysis (EFA) of the main dimensions. The final solution was based on the following criteria: an appropriate number of dimensions consistent with the theoretical background, an explained variance of greater than 50%, and item loadings of greater than 0.4.

Confirmatory Factor Analysis (CFA) within the Structural Equation Modelling (SEM) framework is typically used to evaluate how well a theoretical measurement model fits the observed correlation structure among manifest variables (items). CFA can be applied to validate the factorial structure of an instrument or as the first step in assessing a structural equation model. Since the data are ordinal, a polychoric correlation matrix was used in the CFA, along with varimax rotation, where the Diagonal Weighted Least Squares (DWLS) method was applied to estimate the model parameters. To proceed with the evaluation of model fit, the determinant of the covariance matrix must be different from zero (i.e., the matrix must be positive definite).

The evaluation of model fit is generally performed using (i) fit tests, (ii) goodness-of-fit indices, or (iii) residual analysis and parameter significance (Bentler, 1990). The chi-square (χ^2) goodness-of-fit test theoretically tests whether the fit is perfect. Currently, several indices can be used as alternatives to χ^2, although their empirical basis is sometimes questioned (Steiger et al., 1985). The indices used in this study were chosen based on the most reported fit indices in SEM applications, including χ^2/df (Chi-square divided by degrees of freedom), where a good fit is indicated if <2, acceptable if <5, and unacceptable if >5

(Arbuckle, 2009). The Comparative Fit Index (CFI) adjusts for underestimation in small samples, and CFI < 0.9 indicates poor fit, [0.9–0.95] indicates good fit, and ≥ 0.95 indicates excellent fit (Bentler, 1990). Tucker-Lewis Index (TLI) is similar to CFI, ranges from 0 to 1, and values close to 1 indicate a very good fit.

The Parsimony Goodness-of-Fit Index (PGFI) corrects for artificial model improvement due to additional parameters, and PGFI <0.6 indicates poor fit, [0.6–0.8] indicates moderate fit, and > 0.8 indicates good fit. Root Mean Square Error of Approximation (RMSEA) penalises improvement from simply adding parameters, and RMSEA >0.10 suggests poor fit, values between [0.08–0.10] are mediocre, values between [0.05–0.08] are good, and values <0.05 are very good (Kenny et al., 2015; Steiger et al., 1985).

The statistical analysis was performed using R (version 4.4.3) and the following packages: lavaan, semPlot, and semTools.

3. Results

3.1. Data screening

A preliminary analysis of the data was conducted to identify missing values and inconsistencies. No missing values were found, and the majority of respondents ($n = 101$) used four or five options on the scale, which is expected for a well-distributed questionnaire. The suitability of the data for factor analysis was assessed using Bartlett's Test of Sphericity and the KMO measure of sampling adequacy. Bartlett's Test was statistically significant, $\chi^2_{(56)} = 765.18$, $p < .001$, indicating that the correlation matrix among the items was not an identity matrix. This suggests the presence of significant relationships between the variables, supporting the application of factor analysis. The KMO yielded an overall value of 0.66, indicating moderate suitability for factor analysis (Hutcheson & Sofroniou, 1999). According to the criteria established by Kaiser (1974), values above 0.60 are considered acceptable for factor analysis, but higher values (> 0.80) are preferable. Based on these results, EFA and CFA were deemed appropriate for subsequent analysis.

3.2. EFA

EFA was conducted on the 57 items of the questionnaire using the Minimum Residual (MinRes) method and varimax rotation. The analysis yielded a three-factor solution, explaining 43% of the total variance, which may be considered moderate for a Likert-type scale (ideally, a variance explanation above 50% is recommended). The correlations between the factors and the total scores were high (>0.90), suggesting that the factors adequately represent the item responses. However, as the Kaiser-Meyer-Olkin (KMO) measure of sampling adequacy was below 0.70, it may be necessary to remove variables with low individual adequacy to improve the overall model fit.

3.3. Reliability analysis, factorial validity and convergent validity

From Table 2, a reliability analysis was conducted for the three constructs evaluated in the questionnaire. For the construct DA (27 items), Cronbach's alpha was 0.864, indicating high internal consistency. The item-total statistics revealed that no item markedly improves the internal consistency when deleted, supporting the coherence of the scale. For the DC construct (18 items), the Cronbach's alpha was 0.794, which is also considered acceptable for research purposes. The construct DB (13 items) presented a slightly lower alpha of 0.717, yet still within acceptable limits, particularly for exploratory studies. The global scale yielded a Cronbach's alpha of 0.90, indicating excellent internal consistency across all items.

Regarding factorial validity, the table reports composite reliability (CR) values close to or above the conventional threshold of 0.70 for all constructs (CR = 0.998 for DA and DC; 0.997 for DB), supporting the robustness of the factor structure. These values suggest that the latent constructs are reliably measured by their corresponding items. Furthermore, the Average Variance Extracted (AVE) values were lower than the ideal threshold of 0.50. These values indicate insufficient convergent validity, suggesting that the latent constructs may not be adequately capturing the shared variance among the items.

After item removal, a KMO value of 0.66 was obtained, and Bartlett's test continued to indicate suitability for both EFA and CFA ($\chi^2(45) = 811.56$, $p < .001$). The total variance explained by the three proposed factors increased to 48%.

Table 2
Internal consistency statistics of the 1st proposed scale.

Construct	Cronbach's α	AIIC	CR	AVE
DA (27 items)	0.864	0.181	0.998	0.171
DC (18 items	0.794	0.259	0.998	0.122
DB (12 items)	0.717	0.248	0.997	0.076
Total	**0.9**	**0.1**		

3.4. Discriminant validity

As for discriminant validity, the high inter-construct correlations (Table 3) may indicate limited distinction among constructs. Additionally, when AVE values are lower than the squared correlations between constructs, it raises concerns about the ability of the constructs to discriminate between one another. These results suggest a need for further refinement of the questionnaire.

3.5. Appendix

3.5.1. CFA

Analysing Table 2, several standardised factor loadings (λ_{ij}) exceed 0.50, especially within the DC and DB constructs, reinforcing some evidence of convergent validity at the item level. Nonetheless, a number of loadings fall below the recommended threshold, and some are even negative or very low, pointing to items that may not contribute meaningfully to the construct and could be candidates for removal or revision.

The confirmatory factor analysis (CFA) model fit indices presented in Table 4 suggest that the proposed model does not exhibit an adequate fit to the data based on conventional cut-off criteria. Although the chi-square-to-degrees-of-freedom ratio is below the threshold of 5, which may be considered acceptable, the other fit indices fall short of recommended standards. These findings highlight the need for further model refinement.

3.6. Refinement of the proposed model

The model refinement analysis began with the identification of items presenting factor loadings below 0.40 (Appendix B). A total of 14 such items were identified. However, the initial step involved the removal of items with particularly low factor loadings combined with high residual variances. Accordingly, the following items were excluded from the model: DA_3, DA_23, DA_40, DA_49, DA_51, and DB_20. However, this procedure was still insufficient to achieve a model with satisfactory fit

Table 3
Inter-scale correlations.

	DA	DC	DB
DA	0.413		
DC	0.427	0.350	
DB	0.333	0.913	0.277

Note. Square root of VEM in the principal diagonal.

Table 4
Confirmatory factor analysis model fit indices for the proposed model.

Goodness of fit indices	Reference	Obtained Values
χ^2/df	<5	4.15
CFI	>0.9	0.789
TLI	>0.9	0.781
RMSEA	<0.08	0.173
SRMR	<0.08	0.172

Table 6
Confirmatory factor analysis model fit indices for the final model.

Goodness of fit indices	Reference	Obtained Values
χ^2/df	<5	2.21
CFI	>0.9	0.925
TLI	>0.9	0.921
RMSEA	<0.08	0.107
SRMR	<0.08	0.134

indices, and several items continued to exhibit factor loadings below 0.20. Therefore, modification indices were examined to identify items with substantial cross-loadings across multiple factors. Based on this analysis, the following items were removed one by one from the model: DA_5, DA_12, DA_15, DA_22, DA_31, DA_44, DA_54, DC_41, and DC_23.

The total scale reliability after removing the above-mentioned items was excellent (Cronbach's alpha = 0.90), and the average inter-item correlation for the overall questionnaire was 0.20, supporting internal consistency across all items (Table 5). These results suggest that the refinement process led to a marked improvement in reliability and factorial structure, with acceptable levels of convergent validity across constructs.

The confirmatory factor analysis (CFA) yielded significantly improved fit indices, as shown in Table 6.

The path diagram in Fig. 1 demonstrates that latent factor DA has consistently high loadings across its 16 items, with the majority exceeding 0.70. This indicates a strong relationship between the latent construct and its observed indicators, thereby supporting convergent validity. The lowest loading within this factor is 0.61, still within acceptable limits, and the highest reaches 0.81, which is considered excellent. Similarly, the DC factor also displays adequate to strong standardised loadings, mostly above 0.60, with a maximum of 0.85. These values support the construct's convergent validity, as the items appear to be well-aligned with the latent dimension they represent. One loading is slightly below 0.60 (at 0.56), which is marginal but not uncommon in multidimensional instruments. The third factor, DB, includes 11 items, with loadings ranging approximately from 0.45 to 0.79. Although most loadings exceed 0.60, a few are below this threshold, particularly one item around 0.45. These lower loadings may suggest that some items are less representative of the latent construct, potentially affecting the overall convergent validity of DB. The correlations among the latent factors are also presented in the diagram. The strongest correlation is observed between DB and DC, indicating substantial shared variance and suggesting that these dimensions may not be entirely distinct. The correlations between DA and DC and between DA and DB are moderate, suggesting partial overlap but greater distinctiveness.

4. Discussion

Empathy is widely recognised as a cornerstone of effective psychological practice, underpinning the therapeutic alliance, client engagement, and treatment outcomes (Elliott et al., 2018; Norcross & Lambert, 2018). Promoting empathy not only strengthens therapeutic relationships and enhances patient outcomes, but also contributes to effective and sustainable mental health practice, benefiting both the individuals receiving care and the professionals delivering it (Barrera & de La Motte de Broöns de Vauvert, 2024). However, although the importance of empathy in psychological practice is well documented, most existing instruments focus primarily on cognitive and affective dimensions, often neglecting the behavioural dimension that is essential for a comprehensive assessment (Davis, 1983; Grainger et al., 2023; Lietz et al., 2011). Moreover, most of these measures are administered to students (Bohler et al., 2021; Wilczek-Rużyczka & Kupczyk, 2021) rather than practicing professionals, further limiting their applicability to the complexities of real-world clinical encounters (Gerdes et al., 2011). Addressing this need is crucial, as the accurate assessment of empathy among psychologists is fundamental for ensuring the quality of care provided to clients and for strengthening the therapeutic relationship, which is central to effective psychological interventions (Barrera & de La Motte de Broöns de Vauvert, 2024; Elliott et al., 2018). Therefore, this pilot study sought to address this gap by developing and conducting a preliminary validation of a new self-report instrument designed to assess Portuguese-speaking psychologists' self-perceived empathy in clinical practice.

The results provide initial evidence for the scale's reliability and factorial validity, while also highlighting areas requiring further refinement. Factor analysis was justified by a significant Bartlett's Test and a moderate KMO, aligning with standards for early-stage instrument development. This finding is not uncommon in pilot studies with relatively small samples and newly developed item pools.

The initial EFA yielded a three-factor solution, in line with the three empathy dimensions, explained 43% of the total variance, which, while moderate, is consistent with other Portuguese-language empathy measures (Magalhães et al., 2011; Paro et al., 2012). High inter-factor correlations and moderate KMO suggested some redundancy and the need for item refinement. After systematic exclusion of underperforming items, the final model comprised 43 items across the three factors (DA, DC, DB), with improved psychometric properties. The confirmatory factor analysis (CFA) on the revised scale demonstrated a substantially better fit (CFI = 0.925, TLI = 0.921, χ^2/df = 2.21), supporting the proposed factor structure. Most standardised loadings were above 0.60, especially for DA and DC, indicating strong convergent validity at the item level. DB showed slightly lower loadings for some items, suggesting these may need further revision.

Internal consistency was excellent for the total scale (Cronbach's α = 0.90) and strong for each subscale (α = 0.914, 0.794, and 0.740 for DA, DC, and DB, respectively), indicating that the items within each factor are homogenous and reliably measure the intended constructs. These results are consistent with or exceed values reported for other empathy scales in the literature (Lima & Osório, 2021). Composite reliability was also high. However, AVE values remained below the ideal threshold of 0.50 for some constructs, suggesting room for improvement in convergent validity. Discriminant validity was limited, particularly between the DB and DC factors, as reflected by high inter-construct correlations and AVE values that were lower than the squared correlations. This overlap may reflect conceptual similarities or item redundancy and warrants further theoretical and empirical clarification in future studies. Finally, most items in the refined scale demonstrated satisfactory to strong factor loadings, especially within the DA and DC factors. A few

Table 5
Internal consistency statistics of the final scale.

Construct	Cronbach's α	AIIC	CR	AVE
DA (16 items)	0.914	0.351	0.996	0.459
DC (16 items)	0.794	0.275	0.998	0.413
DB (11 items)	0.740	0.280	0.997	0.393
Total	**0.9**	**0.2**		

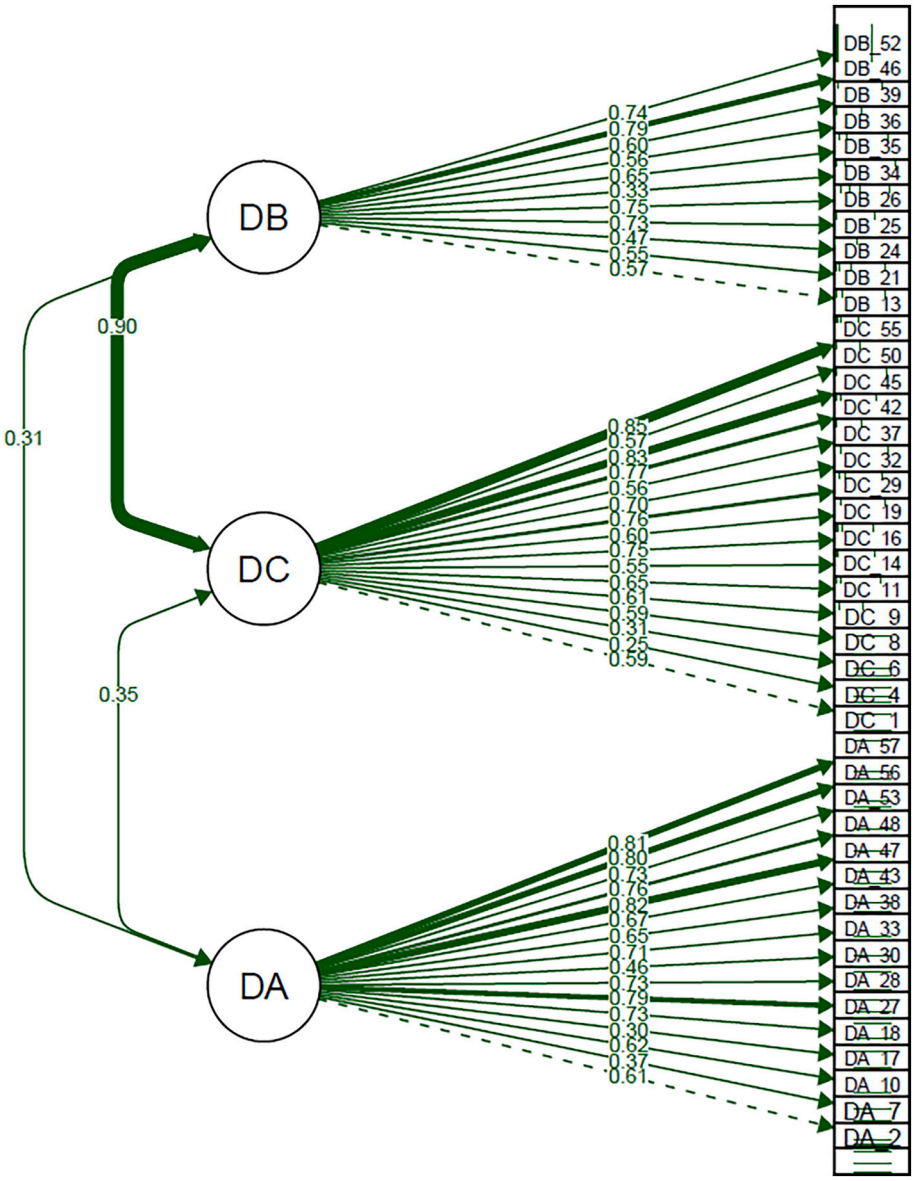


Fig. 1. CFA path diagram. Rectangles represent items, circles represent constructs, values on the arrows are the standardised factor loadings (also referred to as standardised eigenvalues) and curved arrows among latent constructs represent correlations.

items in the DB factor had lower loadings, indicating these may capture more peripheral or less cohesive aspects of empathy as perceived by psychologists.

The findings of this pilot study resonate strongly with existing literature, reinforcing key concepts within the empathy domain. The three-factor structure aligns with the well-established understanding that empathy is a multidimensional construct, comprising cognitive, affective, and behavioural dimensions (Davis, 1983). This structure underscores the complexity of empathy and highlights the ongoing debate regarding the boundaries and interrelations between these facets (Cox et al., 2012; Yamamoto, 2017; Yu & Chou, 2018; Zaki & Ochsner, 2012). The observed factor overlap in this study further echoes this discussion, suggesting that while these three dimensions are distinct, they are not entirely separable in practice (Yamamoto, 2017; Yu & Chou, 2018). Moreover, the moderate explained variance and challenges with discriminant validity reflect broader concerns in the literature regarding the operationalisation and measurement of empathy (Murphy et al., 2020; Stosic et al., 2022).

The findings of this study reinforce the view of empathy as a multidimensional and context-dependent professional competence. The PPES operationalizes empathy through dimensions that mirror the real-world demands of psychological practice, where empathic understanding must be emotionally attuned, cognitively regulated, and behaviourally expressed in therapeutic communication. This approach aligns with contemporary frameworks highlighting the interplay between cognition, emotion, and behaviour in effective therapeutic relationships. From a professional development perspective, these findings emphasize the importance of systematically fostering empathy throughout psychologists' education, supervision, and reflective practice. Training should promote self-awareness, ethical sensitivity, and relational competences. Therefore, a systematic assessment of empathy in psychologists plays a central role in professional development and in the refinement of clinical competencies. By identifying strengths and vulnerabilities across cognitive, affective, and behavioural dimensions, such assessment can inform targeted training strategies and support the acquisition of skills that are empirically linked to therapeutic effectiveness (Elliott et al., 2018; Norcross & Lambert, 2018). Empathy monitoring also promotes critical self-reflection and contributes to more

structured supervision practices, reinforcing ethical sensitivity and relational competence (Gerdes & Segal, 2009; Riker, 2020). Profession-specific instruments further enable the detection of early signs of emotional strain or affective dysregulation and help guide preventive actions that sustain clinicians' well-being (Decety & Jackson, 2004; Murphy & Lilienfeld, 2019). Accordingly, empathy assessment constitutes not only a means of improving the quality of psychological care but also an evidence-based resource for supporting continuous professional growth. The PPES can also serve as a pedagogical tool within ethics training and professional development programmes for psychologists. By promoting reflection on empathic competences across cognitive, affective, and behavioural domains, the scale supports the development of ethical sensitivity and responsible professional practice. This perspective aligns with the principles of University Social Responsibility and the Sustainable Development Goals (SDG), namely SDG 4 – Quality Education, and; SDG 16 – Peace, Justice and Strong Institutions, which highlight the importance of fostering humanistic and socially committed professionals capable of contributing to equitable and sustainable societies.

5. Limitations and future directions

This pilot study provides preliminary evidence for the PPES, although it presents some limitations. First, the same sample was used for both exploratory and confirmatory factor analyses, which may inflate model fit and limit the robustness of the factor structure. Second, although the sample size was adequate for a pilot study, it limits the generalisability and stability of the factor structure (Mundfrom et al., 2005). Third, the moderate KMO and AVE values indicate that convergent validity is not yet optimal. Fourth, high inter-factor correlations may indicate conceptual overlap among components, particularly between cognitive (DC) and behavioural (DB) dimensions of empathy, raise concerns about discriminant validity, suggesting that the current item pool may not fully capture the multidimensional nature of empathy as experienced by psychologists. Five, concurrent validity was not assessed in this phase, comparisons with established measures (e.g., the IRI, JSE, QCAE), and predictive validity against clinical outcomes (e.g., therapeutic alliance, client satisfaction) remain to be conducted. Finally, although we implemented procedural safeguards (anonymity, neutral instructions, randomised item order, balanced keying) as a self-report, the PPES is vulnerable to social desirability and common-method bias. Together, these limitations underscore that the present findings should be interpreted as preliminary and in need of replication and refinement.

Future research should: i) replicate EFA/CFA in larger, independent, and more diverse samples of Portuguese-speaking psychologists (Portugal, Brazil, Angola, Mozambique), to confirm its factor structure and enhance generalisability to the diverse lusophone settings; ii) assess convergent/discriminant validity using HTMT and alternative model forms to probe the DA–DC–DB boundaries, particularly revisiting the conceptual boundaries between cognitive and behavioural dimensions of empathy within psychological practice; iii) assess the predictive and concurrent validity of the PPES by analysing its associations with established measures of clients' perceptions of therapist empathy (Decker et al., 2014) as well as with relevant clinical outcomes, such as patient satisfaction, therapeutic alliance, or treatment effectiveness (Elliott et al., 2018); iv) continue item refinement to increase AVE and improve discriminant validity while preserving theoretical breadth; v) include a brief social desirability scale to quantify response bias and; vi) examine the stability and sensitivity of self-perceived empathy over time through longitudinal studies, which can help determine how empathy evolves in professional development and clinical practice.

To conclude, this study represents an essential initial step in developing a three-component conceptual model of empathy tailored for assessing self-perceived empathy among Portuguese-speaking psychologists. The refined scale shows promising reliability and factorial validity, providing a solid foundation for further refinement. By strengthening the assessment of empathy within psychological practice, this work makes a meaningful contribution to the advancement of empathic skills and supports the broader goal of fostering client-centred care.

CRediT authorship contribution statement

Márcia Cruz: Writing – original draft, Supervision, Project administration, Methodology, Investigation. **Ana Grilo:** Writing – review & editing, Writing – original draft, Software, Methodology, Investigation. **Susana Custódio:** Writing – review & editing, Writing – original draft, Validation, Methodology, Investigation. **Artemisa R. Dores:** Writing – review & editing, Writing – original draft, Validation, Methodology, Investigation. **Carina Silva:** Writing – original draft, Visualization, Validation, Software, Formal analysis. **Paulo Alves:** Writing – original draft, Software, Resources, Methodology, Investigation, Data curation, Conceptualization.

Declaration of competing interest

The authors report there are no competing interests to declare.

Acknowledgments

Our sincere gratitude to all the psychologists who generously dedicated their time to participating in this study. ARD has been supported by Fundação para a Ciência e Tecnologia (FCT) through R&D Units funding (UID/5210/2025). https://doi.org/10.54499/UID/05210/2025.

Appendix A

Construct	Item		Cronbach's α	α if item deleted	AIIC	CR	AVE	λ_{ij}
DA			**0.864**		**0.181**	**0.998**	**0.171**	
(27 items)	DA_2r	I feel uncomfortable when the client shares highly emotional accounts.		0.855				0.583
	DA_3	I feel the urge to protect the client if they are subjected to aggression or humiliation.		0.877				−0.240
	DA_5r	I feel anger towards a client who has committed a crime.		0.856				0.583
	DA_7r	I have felt romantically attracted to a client.		0.863				0.683
	DA_10r	I feel uncomfortable in the presence of a client who has committed a crime.		0.853				0.383
	DA_12r	I feel sadness at the client's suffering.		0.862				0.731
	DA_15r	I feel emotionally upset when I perceive that the client is attempting to be manipulative.		0.851				0.349
	DA_17	I do not suffer from the problems the client is experiencing.		0.865				0.777
	DA_18r	I feel anger when the client shares something that goes against my principles.		0.854				0.330

(*continued on next page*)

(*continued*)

Construct	Item		Cronbach's α	α if item deleted	AIIC	CR	AVE	λ_{ij}
	DA_22r	I feel uncomfortable when the client describes a situation similar to one I have experienced.		0.854				0.727
	DA_27r	I feel embarrassed by some of the client's accounts.		0.852				0.661
	DA_28r	I feel uncomfortable in situations where the client has an intense emotional outburst.		0.853				0.771
	DA_30r	I get emotional when listening to the client's stories.		0.859				0.714
	DA_31	I am sensitive to the feelings that the client conveys beyond their objective discourse.		0.870				0.245
	DA_33r	I feel uneasy in the presence of specific clients.		0.853				0.676
	DA_38r	I feel uneasy when I perceive that the client is attempting to be manipulative.		0.853				0.719
	DA_40	I can sense the client's emotional state even before they describe it.		0.868				0.251
	DA_43r	I feel uncomfortable when the client describes a situation similar to one I am currently experiencing.		0.853				0.707
	DA_44	I feel joy at the client's achievements.		0.866				0.351
	DA_47r	I feel emotional resonance when listening to some of the client's accounts.		0.852				0.780
	DA_48r	I feel uncomfortable with the client's coldness when recounting disturbing situations/ experiences.		0.851				0.738
	DA_49r	The client's misfortunes and problems do not disturb me.		0.880				−0.234
	DA_51	When I perceive that the client is being deceived, I feel the urge to protect them.		0.874				−0.238
	DA_53r	I feel anger or contempt towards hostile clients.		0.852				0.715
	DA_54r	Sometimes, I feel indifferent to or disconnected from the client's emotional state.		0.859				0.595
	DA_56r	I feel uncomfortable when the client shares viewpoints that are different from my own.		0.853				0.795
	DA_57r	I feel ashamed when listening to some of the client's accounts.		0.853				0.774
			0.794		**0.259**	**0.998**	**0.122**	
DC (18 items)	DC_1	I try to understand the client, considering their perspective on situations.		0.788				0.572
	DC_4r	I usually distance myself from the client's perspective when they describe relational issues.		0.809				0.277
	DC_6r	I try to distance myself when the client talks about their problems/concerns, avoiding thinking from their point of view.		0.806				0.316
	DC_8	I find it easy to understand the client's perspective, even when it is clearly different from my own.		0.787				0.601
	DC_9	I try to understand the client's non-verbal language.		0.789				0.603
	DC_11	I try to "put myself in the client's shoes".		0.774				0.607
	DC_14	When the client is suffering, I usually understand how they feel.		0.782				0.525
	DC_16	When the client describes a situation with which I disagree, I try to understand their perspective.		0.772				0.711
	DC_19r	I find it difficult to perceive situations from the client's perspective		0.782				0.633
	DC_23	I usually understand how the client feels, even before they say it.		0.807				0.221
	DC_29	I understand the client's feelings and their reasons for thinking or acting in a certain way.		0.774				0.735
	DC_32	I usually try to put myself in the client's position when they reveal relational issues.		0.771				0.659
	DC_37	I usually notice when the client is in a negative emotional state.		0.789				0.568
	DC_41r	I usually find it difficult to perceive situations from the client's point of view.		0.781				0.753
	DC_42	To understand the client, I pay attention to non-verbal cues and body language.		0.778				0.767
	DC_45	I usually put myself in the client's place to try to understand their feelings and motives.		0.765				0.815
	DC_50	I usually notice when the client is in a positive emotional state.		0.786				0.553
	DC_55	I try to understand the client's feelings.		0.772				0.845
DB			**0.717**		**0.248**	**0.997**	**0.076**	
(12 items)	DB_13r	I tend to express value judgements about the client's opinions and behaviours.		0.692				0.600
	DB_20r	To maintain a professional relationship, I avoid certain social behaviours, such as touch and eye contact.		0.740				0.232
	DB_21	I easily adapt my behaviour to the client's emotional state.		0.690				0.554
	DB_24r	I tend to advise the client about what they should feel or do.		0.703				0.471
	DB_25	I actively listen when the client describes their situation.		0.690				0.733
	DB_26	I offer the client a simplified version of my understanding of what they have expressed.		0.680				0.731
	DB_34	I do not express my opinion when the client describes a situation that goes against my principles and values.		0.720				0.311
	DB_35	I easily adapt my behaviour in situations where the client is very angry.		0.698				0.660
	DB_36	When the client describes a situation like one, I have experienced, even if it affects me, I do not show it.		0.697				0.569
	DB_39	In a situation where the client cries uncontrollably, I accept it and give them space to express their emotions.		0.701				0.598
	DB_46	I can convey to the client my understanding of the emotions/feelings underlying their discourse.		0.692				0.791
	DB_52	I adapt my tone of voice to the client's emotional state.		0.685				0.712
Total			**0.9**		**0.1**			

r: denote reverse items

Appendix B

Items were removed in two passes: first, we excluded items with very low standardised loadings and high residual variances (primary rule: $\lambda < 0.40$ with problematic uniqueness), then iteratively removed items that showed substantial cross-loadings according to modification indices, prioritizing parsimony while preserving theoretical coverage of each domain. All decisions are consistent with the CFA/EFA workflow specified in the manuscript (polychoric matrix, DWLS estimation) and with the thresholds reported for factorial/convergent validity (e.g., λ, AVE, CR). The final scale retained 43 items (DA = 16, DC = 16, DB = 11) with improved model fit (e.g., $\chi^2/df \approx 2.21$, CFI ≈ 0.925, TLI ≈ 0.921) and α ≈ 0.90 for the total score; AVE remained suboptimal in places and discriminant validity was limited, which we acknowledge as preliminary and to be addressed via replication and

refinement.

Table A1
Removed items and justification.

Item	Construct	Statistical trigger	How identified	Theory/semantics	Coverage preserved by…
DA_3	Affective (DA)	Very low loading + high residuals	Pass 1 ($\lambda <$ 0.40)	Content overlapped with stronger DA items; potential wording softness/ambiguity (risk of social-desirability).	Remaining DA set (multiple items with $\lambda \geq 0.60$).
DA_23	DA	Very low loading + high residuals	Pass 1	Redundant nuance relative to retained DA items; limited incremental variance.	DA items retained (α for DA = 0.914).
DA_40	DA	Very low loading + high residuals	Pass 1	Ambiguity/double-barrel risk; weak link to latent DA after estimation.	DA items with $\lambda \geq 0.61$–0.81.
DA_49	DA	Very low loading + high residuals	Pass 1	Overlap with broader affective responsiveness items; weaker psychometrics.	DA core content retained.
DA_51	DA	Very low loading + high residuals	Pass 1	Narrow behavioural phrasing under DA; construct drift risk.	DA/DC boundary handled by stronger items.
DB_20	Behavioural (DB)	Very low loading + high residuals	Pass 1	Item wording likely constrained variance (response set); peripheral to DB core.	DB retained items (several $\lambda \approx$ 0.70+).
DA_5	DA	Cross-loading / modification index	Pass 2 (MI)	Semantics likely straddled affective–cognitive; removed to reduce cross-factor bleed.	DA/DC distinctiveness improved.
DA_12	DA	Cross-loading / modification index	Pass 2	Redundancy with retained DA items; removed for parsimony.	DA α and CR remained high.
DA_15	DA	Cross-loading / modification index	Pass 2	Conceptual overlap; marginal contribution after re-specification.	DA convergent pattern preserved.
DA_22	DA	Cross-loading / modification index	Pass 2	Ambiguity in appraisal vs. felt affect; improved simple structure after removal.	Global fit improved.
DA_31	DA	Cross-loading / modification index	Pass 2	Overlapping nuance with similar DA items; trimmed to avoid multicollinearity.	DA content breadth retained.
DA_44	DA	Cross-loading / modification index	Pass 2	Lexical redundancy; minimised semantic overlap.	DA α remained ≥ 0.91.
DA_54	DA	Cross-loading / modification index	Pass 2	Edge-case affect phrasing; contributed to factor blurring.	DA/DC discriminability marginally improved.
DC_23	Cognitive (DC)	Cross-loading / modification index	Pass 2	Concept drift towards behavioural expression; removed to protect DC coherence.	DC retained items with λ up to $\approx$ 0.85.

Data availability

The datasets generated during and/or analyzed during the current study are not publicly available due a pilot study but are available from the corresponding author on reasonable request.

References

Arbuckle, J. L. (2009). *AMOS 17 user's guide*. SPSS Inc.

Barkham, M., Lutz, W., Lambert, M. J., & Saxon, D. (2017). Therapist effects, effective therapists, and the law of variability. In L. G. Castonguay, & C. E. Hill (Eds.), *How and why are some therapists better than others? Understanding therapist effects* (pp. 13–36). American Psychological Association. https://doi.org/10.1037/0000034-001.

Barrera, A., & de La Motte de Broöns de Vauvert, S. J. G. N. (2024). Empathy and the work of clinical psychiatrists: Narrative review. BJPsych Advances, 30(2), 116–123. doi:https://doi.org/10.1192/bja.2022.80.

Barrett-Lennard, G. T. (1981). The empathy cycle: Refinement of a nuclear concept. *Journal of Counseling Psychology, 28*(2), 91–100. https://doi.org/10.1037/0022-0167.28.2.91

Beaton, D. E., Bombardier, C., Guillemin, F., & Ferraz, M. B. (2000). Guidelines for the process of cross-cultural adaptation of self-report measures. *Spine, 25*(24), 3186–3191. https://doi.org/10.1097/00007632-200012150-00014

Bellosta-Batalla, M., Cebolla, A., Pérez-Blasco, J., & Moya-Albiol, L. (2019). La empatía en el ejercicio y formación de los psicólogos clínicos y sanitaristas, y su relación con mindfulness y compasión. *Revista Argentina de Clínica Psicológica, 28*(2), 210–220. https://doi.org/10.24205/03276716.2019.1095

Benlidayi, I. C. (2024). Translation and cross-cultural adaptation: A critical step in survey research. *Journal of Korean Medical Science, 39*(49), Article e336. https://doi.org/10.3346/jkms.2024.39.e336

Bentler, P. M. (1990). Comparative fit indexes in structural modelling: Adjudging model fit. *Personality and Individual Differences, 42*(5), 815–824.

Bland, M., & Altman, G. (1997). Statistics notes: Cronbach's alpha. *British Medical Journal, 314*, 572.

Bohler, T. E., Brown, R. F., & Dunn, S. (2021). Relationship between affective state and empathy in medical and psychology students. *Australian Psychologist, 56*(4), 311–323. https://doi.org/10.1080/00050067.2021.1926218

Brett, J. D., Becerra, R., Maybery, M. T., & Preece, D. A. (2023). The psychometric assessment of empathy: Development and validation of the Perth empathy scale. *Assessment, 30*(4), 1140–1156. https://doi.org/10.1177/10731911221086987

Cabedo-Peris, J., Martí-Vilar, M., Merino-Soto, C., & Ortiz-Morán, M. (2021). Basic empathy scale: A systematic review and reliability generalization Meta-analysis. *Healthcare (Basel, Switzerland), 10*(1), 29. https://doi.org/10.3390/healthcare10010029

Cai, H., Wu, Q., & Wang, P. (2026). The influence of workplace experience and professional education on enhancing empathy among doctors, nurses and healthcare students: A comparative cross-sectional study. *International Journal of Nursing Studies, 173*, Article 105275. https://doi.org/10.1016/j.ijnurstu.2025.105275

Camões Institute. (2023). The Portuguese language in the world. Home - Camões - Instituto da Cooperação e da Língua.

Child, D. (2006). *The essentials of factor analysis* (3rd ed.). Continuum.

Clark, L. A., & Watson, D. (1995). Constructing validity: Basic issues in objective scale development. *Psychological Assessment, 7*(3), 309–319.

Clark, M. A., Robertson, M. M., & Young, S. (2018). "I feel your pain": A critical review of organizational research on empathy. *Journal of Organizational Behavior, 40*(2), 166–192. https://doi.org/10.1002/job.2348

Clarke, E., DeNora, T., & Vuoskoski, J. (2015). Music, empathy and cultural understanding. *Physics of Life Reviews, 15*, 61–88. https://doi.org/10.1016/j.plrev.2015.09.001

Coutinho, J., Beiramar, A., Silva, C., Lema, A., Lima, V., Grace, R., Oliveira-Silva, P., Gonçalves, Ó., & Sampaio, A. (2015). Evidências de validade da versão portuguesa do Índice de Reatividade Interpessoal para Casais. *Avaliação Psicológica, 14*(3), 309–317. http://pepsic.bvsalud.org/scielo.php?script=sci_arttext&pid=S1677-04712015000300003&lng=pt&tlng=pt.

Cox, C. L., Uddin, L. Q., Di Martino, A., Castellanos, F. X., Milham, M. P., & Kelly, C. (2012). The balance between feeling and knowing: Affective and cognitive empathy are reflected in the brain's intrinsic functional dynamics. *Social Cognitive and Affective Neuroscience, 7*(6), 727–737. https://doi.org/10.1093/scan/nsr051

Cronbach, L. J. (1951). Coefficient alpha and the internal structure of tests. *Psychometrika, 16*(3), 297–334.

Cuff, B. M., Brown, S. J., Taylor, L., & Howat, D. J. (2016). Empathy: A review of the literature. *European Journal of Psychology, 12*(3), 332–342.

Davis, M. H. (1980). A multidimensional approach to individual differences in empathy. *JSAS Catalog of Selected Documents in Psychology, 10*, 85–103.

Davis, M. H. (1983). Measuring individual differences in empathy: Evidence for a multidimensional approach. *Journal of Personality and Social Psychology, 44*(1), 113–126.

Davis, M. H. (1994). *Empathy: A social psychological approach*. Westview Press.

Davis, M. H. (2018). *Empathy: A social psychological approach*. Routledge.

Decety, J., & Cowell, J. M. (2015). Empathy, justice, and moral behavior. *AJOB Neuroscience, 6*(3), 3–14. https://doi.org/10.1080/21507740.2015.1047055

Decety, J., & Jackson, P. L. (2004). The functional architecture of human empathy. *Behavioral and Cognitive Neuroscience Reviews, 3*(2), 71–100. https://doi.org/10.1177/1534582304267187

Decety, J., & Yoder, K. J. (2017). The emerging social neuroscience of empathy: Concepts, controversies, and implications for psychiatry. *Harvard Review of Psychiatry, 25*(5), 227–244. https://doi.org/10.1097/HRP.0000000000000155

Decker, S. E., Nich, C., Carroll, K. M., & Martino, S. (2014). Development of the therapist empathy scale. *Behavioural and Cognitive Psychotherapy, 42*(3), 339–354. https://doi.org/10.1017/S1352465813000039

Elliott, R., Bohart, A., Watson, J., & Murphy, D. (2019). Empathy. In J. C. Norcross, & M. J. Lambert (Eds.), *Psychotherapy relationships that work: Volume 1: Evidence-based therapist contributions* (pp. 245–317). Oxford Academic. https://doi.org/10.1093/med-psych/9780190843953.003.0007.

Elliott, R., Bohart, A. C., Watson, J. C., & Murphy, D. (2018). Therapist empathy and client outcome: An updated meta-analysis. *Psychotherapy, 55*(4), 399–410. https://doi.org/10.1037/pst0000175

Falcone, E. M. de O, Ferreira, M. C., Luz, R. C. M. da, Fernandes, C. S., Faria, C. de A., D'Augustin, J. F., Sardinha, A., & Pinho, V. D. de. (2008). Inventário de Empatia (I. E.): Desenvolvimento e validação de uma medida brasileira. *Avaliação* Psicológica, 7 (3), 321–334. https://pepsic.bvsalud.org/pdf/avp/v7n3/v7n3a06.pdf.

Falk, A., & O'Hara, J. (2021). Empathy: Assessment instruments and psychometric quality. *Frontiers in Psychology, 12*, Article 781346. https://doi.org/10.3389/fpsyg.2021.781346

Fornell, C., & Larcker, D. F. (1981). Evaluation SEM with unobserved variables and measurement error. *Journal of Marketing Research, 18*, 39–50.

Gerdes, K. E., Lietz, C. A., & Segal, E. A. (2011). Measuring empathy in the 21st century: Development of an empathy index rooted in social cognitive neuroscience and social justice. *Social Work Research, 35*(2), 83–93. https://doi.org/10.1093/swr/35.2.83

Gerdes, K. E., & Segal, E. A. (2009). A social work model of empathy. *Advances in Social Work, 10*(2), 114–127. https://doi.org/10.18060/235

Gerdes, K. E., Segal, E. A., & Lietz, C. A. (2010). Conceptualising and measuring empathy. *British Journal of Social Work, 40*, 2326–2343. https://doi.org/10.1093/bjsw/bcq048

Grainger, S. A., McKay, K. T., Riches, J. C., Chander, R. J., Cleary, R., Mather, K. A., … Henry, J. D. (2023). Measuring empathy across the adult lifespan: A comparison of three assessment types. *Assessment, 30*(6), 1870–1883. https://doi.org/10.1177/10731911221127902

Grosseman, S., & Dohms, M. (2021). Comunicação com emoções fortes: Resposta empática à raiva, ao medo e à tristeza no cuidado à saúde. In M. Dohms, & G. Gusso (Eds.), *Comunicação clínica: Aperfeiçoando os encontros em saúde* (pp. 101–113). Artmed Editora.

Gungordu, N., Walker, D. I., & Hernandez-Reif, M. (2025). The role of basic psychological needs and empathy on prosocial behavior in emerging adulthood. *European Journal of Educational Research, 14*(3), 901–915. https://doi.org/10.12973/eu-jer.14.3.901

Habib, H., Niinuma, S. A., Alrefaie, K., et al. (2025). Shades of acceptance and adjustment; a discursive psychological analysis to showcase empathy in medical students. *BMC Medical Education, 25*, 415. https://doi.org/10.1186/s12909-024-06625-2

Hair, J. F., Black, W. C., Babin, B. J., & Anderson, R. E. (2019). *Multivariate Data Analysis* (8th ed.). Cengage Learning.

Hoffman, M. L. (1984). The contribution of empathy to justice and moral judgment. In N. Eisenberg, & J. Strayer (Eds.), *Empathy and its development* (pp. 47–80). Cambridge University Press.

Hogan, R. (1969). Development of an empathy scale. *Journal of Consulting and Clinical Psychology, 33*, 307–316. https://doi.org/10.1037/h0027580

Hojat, M., DeSantis, J., Shannon, S. C., Mortensen, L., Speicher, M., Bragan, L., … Calabrese, L. (2018). The Jefferson scale of empathy: A nationwide study of measurement properties, underlying components, latent variable structure, and national norms in medical students. *Advances in Health Sciences Education, 23*, 899–920. https://doi.org/10.1007/s10459-018-9839-9

Hojat, M., Gonnella, J. S., Nasca, T. J., Mangione, S., Veloski, J. J., & Magee, M. (2002). The Jefferson scale of physician empathy: Further psychometric data and differences by gender and specialty at item level. *Academic Medicine, 77*(10), S58–S60.

Hojat, M., Mangione, S., Nasca, T. J., Cohen, M. J. M., Gonnella, J. S., Erdmann, J. B., … Magee, M. (2001). The Jefferson scale of physician empathy: Development and preliminary psychometric data. *Educational and Psychological Measurement, 61*(2), 349–365. https://doi.org/10.1177/00131640121971158

Hutcheson, G., & Sofroniou, N. (1999). *The multivariate social scientist: Introductory statistics using generalized linear models*. Thousand Oaks, CA: Sage Publication.

Ishtiyaq, S., Kumari, J., Chauhan, R., Srivastwa, H. K., Pandey, A. S., & Tyagi, S. (2024). Relationship between empathy and prosocial behavior. *International journal for. Multidisciplinary Research, 6*(6). https://doi.org/10.36948/ijfmr.2024.v06i06.30999

Kaiser, H. F. (1974). An index of factorial simplicity. Psychometrika, 39(1), 31–36. doi: https://doi.org/10.1007/BF02291575.

Kenny, D. A., Kaniskan, B., & McCoach, D. B. (2015). The performance of RMSEA in models with small degrees of freedom. *Sociological Methods & Research, 44*(3), 486–507. https://doi.org/10.1177/0049124114543236

Konrath, S. H., Meier, B. P., & Bushman, B. J. (2018). Development and validation of the single-item trait empathy scale (SITES). *Journal of Research in Personality, 73*, 111–122.

Konstam, V., Holmes, W., & Levine, B. (2003). Empathy, selfism, and coping as elements of the psychology of forgiveness: A preliminary study. *Counseling and Values, 47*, 172–183. https://doi.org/10.1002/j.2161-007X.2003.tb00264.x

Lee, Y. M., Park, H., Shin, H. S., & Huh, S. (2023). Validation of a modified version of interpersonal reactivity index for medical students using rasch analysis. *Teaching and Learning in Medicine, 35*(3), 346–355. https://doi.org/10.1080/10401334.2022.2042815

Lietz, C. A., Gerdes, K. E., Sun, F., Geiger, J. M., Wagaman, M. A., & Segal, E. A. (2011). The empathy assessment index (EAI): A confirmatory factor analysis of a multidimensional model of empathy. *Journal of the Society for Social Work and Research, 2*(2), 104–124. https://doi.org/10.5243/jsswr.2011.6

Lima, F. F., & Osório, F. (2021). Empathy: Assessment instruments and psychometric quality – A systematic literature review with a Meta-analysis of the past ten years. *Frontiers in Psychology, 12*, Article 781346. https://doi.org/10.3389/fpsyg.2021.781346

Limpo, T., Alves, R. A., & Castro, S. L. (2010). Medir a empatia: Adaptação portuguesa do Índice de Reactividade Interpessoal. *Laboratório de Psicologia, 8*(2), 171–184.

Maddaluno, O., Aiello, E. N., Roncoroni, C., Prunas, A., & Bolognini, N. (2022). The Reading the mind in the eyes test, Iowa gambling task and interpersonal reactivity index: Normative data in an Italian population sample. *Archives of Clinical Neuropsychology: The Official Journal of the National Academy of Neuropsychologists, 37*(5), 929–938. https://doi.org/10.1093/arclin/acab100

Magalhães, E., Costa, P., & Costa, M. J. (2012). Empathy of medical students and personality: Evidence from the five-factor model. *Medical Teacher, 34*(10), 807–812. https://doi.org/10.3109/0142159X.2012.702248

Magalhães, E., Salgueira, A. P., Costa, P., & Costa, M. (2011). Empathy in senior year and first year medical students: A cross-sectional study. *BMC Medical Education, 11*, 52. https://doi.org/10.1186/1472-6920-11-52

Man, X., & Jing, Z. (2025). The role of prosocial tendencies in the relationships between gratitude, perceived social support, and psychological well-being among Chinese university students: A structural equation modeling approach. *Frontiers in Psychology, 16*, Article 1510543. https://doi.org/10.3389/fpsyg.2025.1510543

Manarte, L., & Andrade, A. (2018). Comparative analysis and validation of the Portuguese version of the interpersonal reactivity index. *PsiLOGOS, 16*(1), 45–59. https://doi.org/10.25752/psi.14905

McNeish, D. (2018). Thanks coefficient alpha, We'll take it from Here. *Psychological Methods, 23*(3), 412–433. https://doi.org/10.1037/met0000144

Mehrabian, A., & Epstein, N. (1972). A measure of emotional empathy. *Journal of Personality, 40*(4), 525–543. https://doi.org/10.1111/j.1467-6494.1972.tb00078.x

Mundfrom, D. J., Shaw, D. G., & Ke, T. L. (2005). Minimum sample size recommendations for conducting factor analyses. *International Journal of Testing, 5* (2), 159–168. https://doi.org/10.1207/s15327574ijt0502_4

Muñoz, C. (2020). Función de la empatía en el trabajo de psicólogos y psicoterapeutas. *Revista Scientific*, *(Ed. esp. 5)*, 146–164. doi:10.29394/Scientific.issn.2542-2987.2020.5.E.7.146-164.

Murphy, B. A., Costello, T. H., Watts, A. L., Cheong, Y. F., Berg, J. M., & Lilienfeld, S. O. (2020). Strengths and weaknesses of two empathy measures: A comparison of the measurement precision, construct validity, and incremental validity of two multidimensional indices. *Assessment, 27*(2), 246–260. https://doi.org/10.1177/1073191118777636

Murphy, B. A., & Lilienfeld, S. O. (2019). Are self-report cognitive empathy ratings valid proxies for cognitive empathy ability? Negligible meta-analytic relations with behavioral task performance. *Psychological Assessment, 31*(8), 1062–1072. https://doi.org/10.1037/pas0000732

Norcross, J. C., & Lambert, M. J. (2018). Psychotherapy relationships that work III. *Psychotherapy, 55*(4), 303–315. https://doi.org/10.1037/pst0000193

Paro, H. B., Daud-Gallotti, R. M., Tibério, I. C., Pinto, R. M., & Martins, M. A. (2012). Brazilian version of the Jefferson scale of empathy: Psychometric properties and factor analysis. *BMC Medical Education, 12*, 73. https://doi.org/10.1186/1472-6920-12-73

Peng, L., Jiang, Y., Ye, J., & Xiong, Z. (2024). The impact of empathy on prosocial behavior among college students: The mediating role of moral identity and the moderating role of sense of security. *Behavioral sciences (Basel, Switzerland), 14*(11), 1024. https://doi.org/10.3390/bs14111024

Queirós, A., Fernandes, E., Reniers, R., Sampaio, A., Coutinho, J., & Seara-Cardoso, A. (2018). Psychometric properties of the questionnaire of cognitive and affective empathy in a Portuguese sample. *PLoS One, 13*(6), Article e0197755. https://doi.org/10.1371/journal.pone.0197755

Reniers, R. L., Corcoran, R., Drake, R., Shryane, N. M., & Völlm, B. A. (2011). The QCAE: A questionnaire of cognitive and affective empathy. *Journal of Personality Assessment, 93*(1), 84–95. https://doi.org/10.1080/00223891.2010.528484

Riker, J. H. (2020). Empathy, compassion, and meditation: A vision for a Buddhist self psychology. *Psychoanalytic Inquiry, 40*(5), 327–339. https://doi.org/10.1080/07351690.2020.1766323

Rogers, C. R. (1957). The necessary and sufficient conditions of therapeutic personality change. *Journal of Consulting Psychology, 21*(2), 95–103. https://doi.org/10.1037/h0045357

Sharghi, M. J. (2023). Predicting prosocial personality, other-oriented empathy, helpfulness, and social exchange styles based on value system. *Journal of European Psychology Students, 14*(1), 24–38. https://doi.org/10.59477/jeps.537

Simões, E. D., Trombini-Frick, L., Priolo-Filho, S. R., & Cortez, P. A. (2023). Revisão narrativa de literatura dos instrumentos de empatia utilizados no Brasil e na América Latina. *Revista Brasileira de Terapias Cognitivas, 19*(2), 202–210. https://doi.org/10.5935/1808-5687.20230024

Steiger, J. H., Shapiro, A., & Browne, M. W. (1985). On the multivariate asymptotic distribution of sequential chi-square statistics. *Psychometrika, 50*, 253–263.

Stosic, M. D., Fultz, A. A., Brown, J. A., & Bernieri, F. J. (2022). What is your empathy scale not measuring? The convergent, discriminant, and predictive validity of five empathy scales. *The Journal of Social Psychology, 162*(1), 7–25. https://doi.org/10.1080/00224545.2021.1985417

Voultsos, P., Galanis, P., Dafni, M. A., Velonaki, V. S., Andreou, G. N., & Kovatsi, L. (2024). The Greek Jefferson scale of empathy-medical student version (JSE-S): Psychometric properties and its associated factors. *Behavioral Science, 14*(3), 195. https://doi.org/10.3390/bs14030195

Wilczek-Rużyczka, E., & Kupczyk, M. (2021). Level of empathy and severity of psychopathic traits in medical, law and psychology students. *Acta Neuropsychologica, 19*, 411–426. https://doi.org/10.5604/01.3001.0015.4618

Yamamoto S. (2017). Primate empathy: Three factors and their combinations for empathy-related phenomena. *Wiley interdisciplinary reviews.* Cognitive Science, 8(3), 10.1002/wcs.1431. doi:https://doi.org/10.1002/wcs.1431.

Yu, C., & Chou, T. (2018). A dual route model of empathy: A neurobiological prospective. *Frontiers in Psychology, 9*, 2212. https://doi.org/10.3389/fpsyg.2018.02212

Zaki, J., & Ochsner, K. N. (2012). The neuroscience of empathy: Progress, pitfalls and promise. *Nature Neuroscience, 15*(5), 675–680. https://doi.org/10.1038/nn.3085