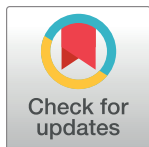


## RESEARCH ARTICLE

# The Spanish version of the reflective functioning questionnaire: Validity data in the general population and individuals with personality disorders

Eduardo Ruiz-Parra<sup>1,2,3\*</sup>, Guadalupe Manzano-García<sup>4</sup>, Roberto Mediavilla<sup>3,5,6</sup>, Beatriz Rodríguez-Vega<sup>5,6,7</sup>, Guillermo Lahera<sup>8,9</sup>, Ana I. Moreno-Pérez<sup>8,9</sup>, Alberto M. Torres-Cantero<sup>10,11</sup>, Juan Rodado-Martínez<sup>12,13</sup>, Amaia Bilbao<sup>14,15,16</sup>, Miguel Ángel González-Torres<sup>1,2,3</sup>



## OPEN ACCESS

**Citation:** Ruiz-Parra E, Manzano-García G, Mediavilla R, Rodríguez-Vega B, Lahera G, Moreno-Pérez AI, et al. (2023) The Spanish version of the reflective functioning questionnaire: Validity data in the general population and individuals with personality disorders. PLoS ONE 18(4): e0274378. <https://doi.org/10.1371/journal.pone.0274378>

**Editor:** Sascha Müller, Universität Kassel: Universität Kassel, GERMANY

**Received:** May 17, 2021

**Accepted:** August 26, 2022

**Published:** April 6, 2023

**Copyright:** © 2023 Ruiz-Parra et al. This is an open access article distributed under the terms of the [Creative Commons Attribution License](https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

**Data Availability Statement:** All Spanish RFQ validation datafiles are available as "S1 Dataset" (as an.xlsx file) in [Supporting Information](#). Identification, socio-demographic and diagnostic standardisation data are not accessible to protect the confidentiality of participants.

**Funding:** The study received a Health Strategic Action fund from the Carlos III Health Institute of the Spanish Ministry of Science and Innovation. ID P116/02058. <https://www.isciii.es/QueHacemos/>

**1** Department of Neurosciences, Faculty of Medicine and Nursing, University of the Basque Country, Bilbao, Spain, **2** Department of Psychiatry, Basurto University Hospital, Osakidetza Basque Health Service, Bilbao, Spain, **3** Biomedical Research Networking Centre in Mental Health (CIBERSAM), Madrid, Spain, **4** Department of Educational Sciences, University of La Rioja, Logroño, Spain, **5** Department of Psychiatry, School of Medicine, Autonomous University of Madrid (UAM), Madrid, Spain, **6** La Paz Hospital Institute for Health Research (IdiPAZ), Madrid, Spain, **7** Department of Psychiatry, Clinical Psychology and Mental Health, La Paz University Hospital, Madrid, Spain, **8** Department of Medicine and Medical Specialties, Faculty of Medicine and Health Sciences, University of Alcalá, Alcalá de Henares, Madrid, Spain, **9** Department of Psychiatry, Príncipe de Asturias University Hospital, Alcalá de Henares, Madrid, Spain, **10** Department of Public Health Sciences, School of Medicine, University of Murcia, Murcia, Spain, **11** Department of Preventive Medicine, Virgen de la Arrixaca University Clinical Hospital, El Palmar, Murcia, Spain, **12** Department of Psychiatry, School of Medicine, University of Murcia, Murcia, Spain, **13** Department of Psychiatry, Reina Sofía University Hospital, Murcia, Spain, **14** Research Unit, Basurto University Hospital, Osakidetza Basque Health Service, Bilbao, Bizkaia, Spain, **15** Health Service Research Network on Chronic Diseases (REDISSEC), Madrid, Spain, **16** Kronikgune Institute for Health Services Research, Barakaldo, Spain

\* [eduardo.ruizparra@osakidetza.eus](mailto:eduardo.ruizparra@osakidetza.eus)

## Abstract

### Introduction

Mentalization or reflective functioning (RF) is the capacity to interpret oneself or the others in terms of internal mental states. Its failures have been linked to several mental disorders and interventions improving RF have a therapeutic effect. Mentalizing capacity of the parents influences the children's attachment. The Reflective Functioning Questionnaire (RFQ-8) is a widely used tool for the assessment of RF. No instrument is available to assess general RF in Spanish-speaking samples. The aim of this study is to develop a Spanish version of the RFQ-8 and to evaluate its reliability and validity in the general population and in individuals with personality disorders.

### Methods

602 non-clinical and 41 personality disordered participants completed a Spanish translation of the RFQ and a battery of self-reported questionnaires assessing several RF related constructs (alexithymia, perspective taking, identity diffusion and mindfulness),

[Financiacion/solicitudes/Paginas/default.aspx](#) No researcher have received money from the fund. This kind of public funds are managed by public institutes (in this case BIOEF) that ensure this money is only used for some of the material resources necessary to carry out the investigation. <https://www.bioef.org/es/> The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

**Competing interests:** The authors have declared that no competing interest exist.

psychopathology (general and specific) and interpersonal problems. Temporal stability was tested in a non-clinical sub-sample of 113 participants.

## Results

Exploratory and confirmatory factor analyses suggested a one-factor structure in the Spanish version of the RFQ-8. RFQ-8 understood as a single scale was tested, with low scorings reflecting genuine mentalizing, and high scorings uncertainty. The questionnaire showed good internal consistence in both samples and moderate temporal stability in non-clinical sample. RFQ correlated significantly with identity diffusion, alexithymia, and general psychopathology in both samples; and with mindfulness, perspective taking, and interpersonal problems in clinical sample. Mean values of the scale were significantly higher in the clinical group.

## Discussion

This study provides evidence that the Spanish version of the RFQ-8, understood as a single scale, has an adequate reliability and validity assessing failures in reflective functioning (i.e., hypomentalization) in general population and personality disorders.

## Introduction

Mentalization is the capacity to understand ourselves or the others in terms of mental states (e.g., feelings, wishes, goals, desires and attitudes) [1]. As a result thereof, behavior and emotional experiences become more meaningful and predictable, especially in the context of close and intimate relationships [2].

Several dimensions of mentalization have been described, including explicit or automatic versus implicit or controlled mentalization; self-focused versus other-focused mentalization; mentalization based on internal experience versus mentalization based on external cues; and cognitively versus affectively focused mentalization [2].

Genuine mentalization combines (a) the ability to form relatively accurate models of the mind, with (b) the awareness that any certainty is conditioned by the inevitable opacity of mental states [1]. On the basis of these two aspects, two types of failures in mentalization have been described: hypomentalization, reflecting excessive uncertainty about mental states; and hypermentalization, reflecting excessive certainty about mental states in the absence of appropriate evidence [1, 3].

Failures in mentalization have been linked to a vulnerability to various psychopathological conditions [4], and the interventions aimed at improving mentalization have proven to positively modify its course [5]. Furthermore, it has been suggested that the ability of the caregivers to mentalize in their relationship with the infant may determine the security of the child's attachment [6].

Mentalization has been related to constructs such as theory of mind, mindreading, social cognition, metacognition, empathy, mindfulness, alexithymia, identity, and others. They are not completely equivalent concepts, but valid and easy-to-use measurement instruments are available for them. These instruments have been used as indirect measures of mentalization. However, they only measure some dimensions of mentalization and not others (e.g., measuring alexithymia with the well-known Toronto Alexithymia Scale can only indirectly assess

oneself and emotional dimensions of mentalization), and depart from other theoretical premises, so their use is limited [7, 8].

Reflective function (RF) is an operationalization for research purposes of the mental capacities that generate mentalization [6]. In order to measure them, the Reflective Functioning Scale (RFS) [9], which evaluates RF on transcriptions of the Adult Attachment Interview (AAI) [10], was developed. This instrument constitutes the gold standard for RF measurement. Subsequently, other instruments were designed to measure RF and parental reflective functioning (PRF), focused on the representations that parents have about their own children, reproducing the narrative coding system of the RFS [11–16]. More recently, the Computerized Reflective Functioning (CRF) [17] has been developed. It identifies linguistic markers associated with high RF in narrative transcripts, and can be applied under induced psychological stressful situations [18], thereby facilitating the assessment of automatic mentalization [19].

As RF research progressed, it became apparent that narrative coding instruments were complex and required substantial resources in terms of training and time, resulting in difficulties in conducting large sample studies [20]. The need for easy-to-apply and easy-to-score questionnaires arose. Accordingly, some instruments have been developed: (a) the Mentalizing Stories for Adolescents (MSA) [21], assessing RF in adolescents; (b) the Parental Reflective Functioning Questionnaire (PRFQ) [22], assessing PRF; and, (c) assessing RF in adults, the Mentalization Questionnaire (MZQ) [23], the Mentalization Scale (MentS) [24], the Certainty About Mental States Questionnaire (CAMSQ) [25], the Multidimensional Mentalizing Questionnaire (MMQ) [26], and the 8-item Reflective Functioning Questionnaire (RFQ-8) [27], which also has a 6-item version [28, 29], a 15-item version (RFQ-15) [30], and a 18-item version (RFQ-18) [31]. Some modified versions of the RFQ have been developed for adolescents: the 8-item Reflective Functioning Questionnaire for Youth (RFQY) [4], the 23-item PFQY Scale B [32], and the 6-item RFQY Scale B [32]. The RFQ8 was developed by the group that coined the term mentalization from a more extensive questionnaire, the RFQ-54 [33], whose psychometric assessment data have not been published.

The RFQ-8 is a self-report questionnaire originally developed in English, which has been translated into several languages [34]. Its psychometric properties have been evaluated in general population in English, French, Italian, Greek, German, Persian and Polish; in individuals with personality disorder, and other psychiatric patients, in English, Italian, and German; and in adult type 1 diabetes population in Greek [4, 27–29, 35–38].

The RFQ-8, as originally conceived, contains two subscales, the certainty about mental states scale (RFQc) (items 1 to 6), and the uncertainty about mental states scale (RFQu) (items 2, 4, 5, 6, 7 and 8). High scorings of the scales are assumed to reflect respectively hypermentalization and hypomentalization [27].

In developing the RFQ, two types of items were designed in relation to their scoring system: some used a median-scoring method (extreme answers reflected lower scores, while responses reflecting genuine mentalization—i.e., less extreme answers—received the highest scores), and the others, the so called Scale B items [34], used a polar-scoring method (stronger agreement—or disagreement in case of inverted items—yielded higher RF scores). Only median-scored items were considered for further development of the questionnaire, but, as extreme answers for these items did not discriminate between hypermentalization and hypomentalization, the decision was made to convert them into polar-scored items, using a double-scoring system (scoring them in one direction to reflect hypomentalization, and in the opposite to reflect hypermentalization). For example, the response to the item “I always know what I feel” were rescored from 1, 2, 3, 3, 2, 1 to 1, 2, 3, 4, 5, 6 to account for certainty scale, and to 6, 5, 4, 3, 2, 1 to account for uncertainty scale. In order to capture more extreme levels of the variables, the items were then rescored as 1 = 0, 2 = 0, 3 = 0, 4 = 0, 5 = 1, 6 = 2 and 1 = 2, 2 = 1, 3 = 0, 4 = 0,

5 = 0, 6 = 0. Additionally, the initial 6-point Likert-type scale was changed to a 7-point Likert-type one, rescaling the items from 2, 1, 0, 0, 0, 0 to 3, 2, 1, 0, 0, 0, and from 0, 0, 0, 1, 2 to 0, 0, 0, 1, 2, 3. The final 6 items of each scale of the RFQ-8 were selected by the criteria of showing the highest loadings on their respective factor across a series of exploratory and confirmatory factor analyses (CFA) [27]. Since items 2, 4, 5 and 6 belong to both scales, they are double-scored.

The RFQ-8 has been extensively used to assess RF in research aimed at (a) defining its dysfunction (usually hypomentalization) in groups of patients or other populations [39–47], and how it mediates symptomatic expression [35, 48–67]; (b) assessing change in psychotherapy [68–75]; and (c) understanding the relationship of RF to attachment and parenting [27, 60, 76–79]. Surprisingly, much of this research [27, 43–48, 50, 52, 53, 57, 59–68, 71, 77] suggest that higher RFQc scores may indicate more appropriate mentalization [51, 68], perhaps reflecting confidence in mental states as valid explanations for emotional experiences and behaviors [48].

Recently, during German validation studies [28, 29], some concerns have arisen about the double-scoring procedure, the factor structure of the RFQ, and, again, the validity of RFQc scale. Double-scoring causes problems in factor analysis. Given that only one rating is provided for four of the eight items on the 7-point scale, eight rescaled scores on RFQc and RFQu are mutually determined. Those pairs of scores are not independent and information overlaps. Nine of the 16 theoretical combinations for each of the two paired scores are mathematically impossible. This results in polychoric correlations between several scores approaching  $r = -1$ , clearly indicating that the double-scores are redundant [29]. The two-factor model, as proposed in initial validation studies [4, 27, 35] has been questioned arguing that a negative correlation between the two factors is artificially induced because the residual correlations of double-scored item pairs are restricted to zero [29]. A change in scoring procedure has been proposed, avoiding double-scoring and rescuing the scoring previously used in the RFQ design process (i.e., 1, 2, 3, 4, 5, 6, 7), except for item 7, reversely scored due to its content polarity with the other items [28, 29]. Additionally, evidence has been provided via exploratory factor analysis (EFA) and CFA that a one-factor model can sufficiently explain the observed covariation of the responses to RFQ-8 items when using this new way of scoring [28, 29, 38]. Finally, neither RFQ-8 certainty pole (using new proposed way of scoring), nor its certainty scale (using the previously proposed way of scoring) tend to show positive associations with negative outcomes (i.e., psychopathology) when looking for U shape [28, 29] or linear correlations [27, 38, 43–48, 50, 52, 53, 57, 59–68, 71, 77], thus suggesting the inability of the RFQ-8 to assess hypermentalization [29, 51]. In this sense, the CAMSQ [25] has been recently validated to fill the gap of specifically assess hypermentalization.

As no instruments assessing general RF in Spanish are available (only those assessing PRF), the need for reliable and valid instruments to measure RF in Spanish has been pointed out [8]. The RFQ-8, provided that the issues associated with its scoring system and factor structure are properly analyzed, could be a reasonable option to address this need, since it is an easy-to-use instrument, designed to measure mentalization difficulties, and has already been widely used in research.

The purpose of the present study is to develop a Spanish version of the RFQ-8 conceptually, semantically and operationally equivalent to the original version, and to assess its reliability and construct validity in two samples of Spanish general population and patients with personality disorder.

## Material and methods

### Participants

A non-clinical convenience sample drawn from general adult population was selected from different Spanish Autonomous Communities (Basque Country, La Rioja, Murcia, and Madrid). The sample was recruited by means of informative talks to pre and postgraduate students of a medical school and a business school, and, by word of mouth, among hospital and university staff, their families and friends. Inclusion criteria were to be over 18 years old, and to provide written informed consent. Exclusion criteria were showing general verbal or Spanish language communication difficulties, a clinical diagnosis of cognitive impairment, or having undergone psychiatric or psychological treatment on a mental health facility during the year prior to inclusion in the study. The final sample included 602 participants. A sub-sample of 113 participants was selected by convenience to evaluate retest reliability.

A clinical convenience sample (from Basque Country and Madrid) of 41 adult participants with a diagnosis of personality disorder assessed by means of the SCID-II interview [80] was selected in order to assess the psychometric properties of the questionnaire in a clinical sample.

The study protocol was approved by the Basque Country Clinical Research Ethics Committee.

### Translation of the questionnaire

The original English RFQ was translated into Spanish using a standard translation back-translation procedure aimed at ensuring conceptual, semantic and operational equivalence [81]. Translation was carried out by independent native Spanish and English translators [82]. Clarity, appropriateness and equivalence of the questionnaire were assessed by a group of experts in psychometric evaluation. In order to assess comprehension, a pilot study with 10 subjects with a maximum academic level of secondary education or equivalent (compulsory in Spain) was also included. The original and the Spanish versions of the questionnaire have been included with in S1 and S2 Appendices.

### Measurements

Sociodemographic data, and diagnostic standardization data in the clinical sample, were obtained at the university and clinical facilities linked to the study. Questionnaires were completed online at home in the non-clinical sample, and via paper and pencil, and on line at the mentioned facilities in the clinical sample. Retest were completed 2–3 weeks after first test.

The subjects completed a battery of questionnaires, which included:

The Spanish version of the 8-item **Reflective Functioning Questionnaire** (RFQ-8): each item of the questionnaire is scored on a 7-point Likert-type scale, ranging from "strongly disagree" to "strongly agree". Two different ways of scoring are used, (a) one reflecting the originally proposed for the two scales [27], where the RFQc items are rescored as 1 = 3, 2 = 2, 3 = 1, 4 = 0, 5 = 0, 6 = 0, 7 = 0; RFQu items are rescored as 1 = 0, 2 = 0, 3 = 0, 4 = 0, 5 = 1, 6 = 2, 7 = 3, except for item 7, which is rescored as 1 = 3, 2 = 2, 3 = 1, 4 = 0, 5 = 0, 6 = 0, 7 = 0; and each scale final score is the average of its item scores [34]; (b) and the other, reflecting a single scale [28, 29, 38], keeping all the items in the scoring previously used in the RFQ design process (scoring 1 to 7, and inversely rescored for item 7). The scale final score is the average of all the items scores, with high values indicating uncertainty about mental states and, as previous research suggests [28, 29, 38], low values indicating genuine mentalization.

The **Perspective Taking** subscale of the **Interpersonal Reactivity Index (IRI)** [83]: the IRI is a 28-item self-report scale that assesses implicit empathy. It has been validated in the Spanish general population [84]. The IRI has four subscales: Perspective Taking (PT), Fantasy, Empathic Concern, and Interpersonal Stress. PT include 7 items, and indicates the spontaneous tendency to adopt the psychological point of view of the others (a cognitive dimension of empathy). Each item is scored on a 5-point scale anchored by 0 (A = "does not describe me well") and 4 (D = "describes me very well"). A weak correlation ( $r = -0.18, p < 0.05$ ;  $r = -0.26, p < 0.01$ ) has been reported between PT and RFQu [27, 37]. Internal consistency of the scale in the present study in non-clinical sample ( $\alpha = 0.82$ ) and clinical sample ( $\alpha = 0.92$ ) was good.

The **Toronto Alexithymia Scale (TAS-20)** [85, 86] is a 20-item self-report scale which has been validated both in the Spanish general population and in psychosomatic patients in Spain [87]. The TAS-20 consists of a total alexithymia score (TAS<sub>t</sub>), and three subscales: Difficulty Identifying Feelings (DIF), Difficulty Describing Feelings (DDF), and, less relevant for our study, Externally Oriented Thinking. Each item is evaluated according to a Likert-type scale from 1 to 6, ranging from "very much in disagreement" to "very much in agreement", and rescored to 0 to 5. Both scales of the RFQ showed moderate correlations with TAS<sub>t</sub> (RFQc  $r = -0.53, p < 0.05$ ; RFQu  $r = 0.43, p < 0.05$ ), and DIF subscale (RFQc  $r = -0.45, p = 0.001$ ; RFQu  $r = 0.57, p = 0.001$ ; RFQc  $r = -0.56, p < 0.01$ ; RFQu  $r = 0.50, p < 0.01$ ) in previous validation studies [4, 35, 37]. Internal consistency in our non-clinical sample (TAS<sub>t</sub>  $\alpha = 0.87$ , DIF  $\alpha = 0.87$ , DDF  $\alpha = 0.85$ ) and clinical sample (TAS<sub>t</sub>  $\alpha = 0.89$ , DIF  $\alpha = 0.93$ , DDF  $\alpha = 0.75$ ) was good.

The Spanish version of the **Mindful Attention Awareness Scale (MAAS)** [88, 89] is a 15-item self-report scale that assesses an individual's dispositional capacity to be attentive and aware of the experience of the present moment in daily life (mindfulness). The questionnaire has been validated in Spanish clinical and general populations [90]. It is scored with a Likert-type scale with a range from 1 ("almost always") to 6 ("almost never"). Weak correlations between MAAS and RFQu scale ( $r = -0.33, p < 0.01$ ;  $r = -0.24, p < 0.01$ ), and weak or no correlation with RFQc scale ( $r = 0.34, p < 0.01$ ;  $r = 0.08, p: n.s.$ ) have been reported [27, 37]. Internal consistency in the present non-clinical sample ( $\alpha = 0.90$ ) and clinical sample ( $\alpha = 0.92$ ) was excellent.

The 83-item **Personality Organization Inventory (IPO-83)** [91]: It's a 83-item self-report inventory, validated in Spanish general and clinical population [92], that evaluate the dimensions of the structural organization of the personality. The IPO-83 is composed of (a) three primary scales: Identity Diffusion (21 items), use of Primitive Defenses (16 items), and Reality Testing (20 items); and (b) two secondary scales: Aggression (18 items), and Moral Values (11 items, 3 of them are shared with the primary scales). The items are evaluated on a 5-point Likert scale, ranging from "never true" to "always true" [91]. The primary scales of a larger version, the IPO-136, showed moderate correlations with the RFQu in the original validation study (ID  $r = 0.57, p < 0.01$ ; PD = 0.52,  $p < 0.01$ , RT = 0.54,  $p < 0.01$ ), but only ID scale moderately correlated with the RFQc ( $r = -0.41, p < 0.01$ ) [27]. The total scores of the IPO-16, a shorter version of the inventory—composed only of the primary scales—, used as indicators of severity of personality disfunction, correlated with the unidimensional RFQ-8 ( $r = 0.72, p < 0.001$ ;  $r = 0.64, p < 0.05$ ) [29]. Internal consistency of all the scales in the present non-clinical sample ( $\alpha$  from 0.79 to 0.92), and clinical sample ( $\alpha$  from 0.73 to 0.94) was good.

The **Symptom Checklist 90 Revised (SCL-90-R)** [93]: it's a widely used 90-item self-report checklist, validated in Spanish general population [94, 95]. It assesses 9 symptom patterns scales (Somatization, Obsessive-compulsive, Interpersonal Sensitivity, Depression, Anxiety, Hostility, Phobic Anxiety, Paranoid Ideation, and Psychoticism scales) and 3 psychological discomfort indexes (Global Severity Index -GSI-, Positive Symptom Total -PST-, and Positive



Symptom Distress Index -PSDI-). It is scored with a 5-point Likert-type scale with a range from 0 ("absence of the symptom") to 4 ("total presence of the symptom"). Psychological distress, as measured by a short version of the checklist, the SCL-10, moderately correlated with RFQc ( $r = -0.47$ ,  $p < 0.01$ ) and RFQu ( $r = 0.59$ ,  $p < 0.01$ ) in a sample of diabetic patients [36]. Internal consistency of all the scales in the present non-clinical sample ( $\alpha$  from 0.80 to 0.98), and clinical sample ( $\alpha$  from 0.86 to 0.99) was good.

The **Beck Depression Inventory-II** (BDI-II) [96]: it's a 21-item self-report scale that assesses the presence and severity of depressive symptoms. It has been validated in Spanish general and clinical populations [97–99]. Items are scored in a 4-point Likert-type scale, ranging from total absence of the symptom to most severe presence of it; except for items 16 and 18 that scored in a 7-point scale, but are conveniently rescored to 0 to 4. Moderate correlations of the BDI-II with RFQu in clinical ( $r = 0.53$ ,  $p < 0.01$ ) and non-clinical samples ( $r = 0.40$ ,  $p < 0.01$ ) have been reported [27]. Internal consistency of the inventory in the present non-clinical sample ( $\alpha = 0.92$ ) and clinical sample ( $\alpha = 0.93$ ) was excellent.

The **Personality Inventory for DSM-5 Brief Form** (PID-5-BF) [100, 101]: it's a 25-item self-report scale, validated in Spanish general and clinical populations [102], which evaluates 5 domains of personality traits (Negative Affect, Detachment, Antagonism, Disinhibition and Psychoticism), dysfunctional variants of the Big Five model [103]. Items are scored on a 4-point scale, ranging from 0 ("totally false or often false") to 3 ("very true or often true"). RFQ-8 showed moderate correlations with PID-5-BF total score ( $r = 0.67$ ,  $p < 0.001$ ); and with Negative Affect ( $r = 0.60$ ,  $p < 0.001$ ), Detachment ( $r = 0.46$ ,  $p < 0.001$ ), Disinhibition ( $r = 0.44$ ,  $p < 0.001$ ), and Psychoticism ( $r = 0.52$ ,  $p < 0.001$ ) domains in one study [29]. Internal consistency of PID-5-BF total score in our non-clinical sample ( $\alpha = 0.87$ ) and clinical sample ( $\alpha = 0.92$ ) was good.

The PID-5-BF was designed as a tool to support the mixed categorical-dimensional alternative model of personality disorders in the DSM-5. This model requires assessment on a continuum of personal (identity and self-direction) and interpersonal (empathy and intimacy) functioning levels, and combines it with the assessment of 5 major personality domains [101]. In the present study, in order to assess the relationships between failures in mentalization and the different dimensions of personality, the instrument is used independently from the assessment of personality functioning.

The 32-item **Inventory of Interpersonal Problems** (IIP-32) [104]: it's a self-report inventory that assesses interpersonal functioning, validated in Spanish general population and individuals with personality disorders [105]. It assesses interpersonal behaviors that the subject has difficulty carrying out or carries out excessively using a 5-point Likert-type scale, ranging from 0 ("not at all") to 4 ("extremely"). It provides an overall score, the one used in our study, and scores on eight scales that reflect different interpersonal dimensions. Using the IIP-62, a larger version of the inventory, weak correlations between a general measure of interpersonal problems and RFQc ( $r = -0.16$ ,  $p < 0.05$ ), and RFQu ( $r = 0.32$ ,  $p < 0.01$ ) have been reported [27]. The IIP-32 overall score moderately correlated with the RFQ-8 ( $r = 0.54$ ,  $p < 0.001$ ) in another study [29]. Internal consistency of the overall score in the present non-clinical sample ( $\alpha = 0.90$ ) and clinical sample ( $\alpha = 0.92$ ) was excellent.

## Data analysis

Given the problems with the double-scoring procedure, which reproduce with our data (S1 Table); and the growing evidence towards a lack of specificity of the certainty scale, and a one-factor model, the decision was made to use the new way of scoring in the present validation study. Nevertheless, all analysis were performed using both ways of scoring. The

results using the originally proposed way of scoring [27] can be found in the Supporting Information section.

The percentage of subjects scoring at the lowest possible level of the scale (floor effect) and the highest possible level (ceiling effect) were examined. Floor and ceiling effects should be minimal, and we used 15% as the critical value for those effects [106].

**Construct validity: Factor analyses.**

To study the structural validity of the questionnaire, replicating the methodology used in by Spitzer et al. [28] and Müller et al. [29], both confirmatory factor analysis (CFA) and exploratory factor analysis (EFA) were performed on the non-clinical sample using the new scoring system. CFA was used to investigate the hypothesized, recently proposed, one-factor structure [28, 29]; and the two-factor structure of the RFQ as proposed by the creators of the questionnaire [27]. Weighted least squares mean and variance adjusted (WLSMV) estimator was used. Different fit indices were evaluated using the following criteria: (a) chi squared divided by the degrees of freedom, the result of which had to be  $\leq 3$  to be acceptable; (b) the root mean squared error of approximation (RMSEA), where a value  $< 0.08$  was considered acceptable; and (c) the Tucker-Lewis Index (TLI) and the comparative fit index (CFI), both of which had to be  $> 0.90$  to be satisfactory [27, 107–110]. Factor loadings were also examined, and those  $\geq 0.30$  were considered acceptable [111]. Therefore, if the model surpassed these acceptability criteria, it was considered acceptable. The Lagrange multiplier test, which identifies paths or covariances that should possibly be added to the model to improve the fit was used when the model needed modification. For the comparison of the two-factor and one-factor models, as the models are non-nested, we used the Akaike information criterion (AIC) and the Bayesian information criterion (BIC), where lower values indicate that the model fits better [112]. Further, the two-factor EFA was performed considering the Promax oblique factor rotation. Factor loadings were also examined, and those  $\geq 0.30$  were considered acceptable [111].

Additionally, a CFA was performed using the originally proposed way of scoring [27], based on the creators' recommendations [34], and using the same goodness-of-fit indices.

**Reliability.**

Internal consistency of the RFQ-8 subscales was assessed with Cronbach's alpha coefficient [113]. A coefficient over 0.70 was considered acceptable [114]. The temporal stability was examined by performing a test-retest and calculating the intraclass correlation coefficient (ICC). Values from 0.50 to 0.75 pointed to moderate reliability, from 0.75 to 0.90 to good reliability, and values over 0.90 indicated excellent reliability [115, 116].

**Construct validity: Convergent/discriminant and known-groups validity.**

We assessed convergent and discriminant validity by analyzing the relationship between the RFQ-8 and identity diffusion, perspective taking, mindfulness, and alexithymia with Spearman correlation coefficients. We hypothesized a significant moderate to strong correlation (0.40 to 0.79 in absolute value) [117] between the RFQ-8 scale and the related constructs.

The relationship between the RFQ-8 domains and different measures of psychopathology (SCL-90, BID-II), personality domains (PID-5-BF) and interpersonal functioning (IIP-32) was assessed, again using the Spearman correlation coefficient.

We examined known-groups validity by comparing non-clinical with clinical group. We hypothesized that the comparison will significantly discriminate between both groups. For the comparison, the t-test was used, or the non-parametric Wilcoxon or Kruskal-Wallis test when normality was not met.

Effects were considered significant at  $p < 0.05$ . Statistical analyses were performed with SAS<sup>®</sup> for Windows statistical software, version 9.4 (SAS Institute, Inc., Carey, NC); and MPlus, version 6.1 [118].



## Results

### Demographics

Both the non-clinical and the clinical sample were predominantly females (respectively 82.56%, and 78.05%). Medium age of subjects was 24.28 years (SD = 10.32) in non-clinical sample and 40.44 years (SD = 10.66) in clinical sample. Educational level was higher for non-clinical sample, with only 2.16% with a maximum level of secondary school. As a large part of the sample was recruited among university students, most of the sample (74.42%) had a high school education. The educational level of the clinical sample was more balanced, with 34.5% of subjects with a maximum level of secondary education, and 29.25% with university education, the latter probably in relation to the age of the subjects

Neither ceiling effect, nor floor effect was noted for the questionnaire using the new scoring method, contrasting with a slight floor effect found for the RFQu scale (15.45% of the subjects scored at the lowest possible level of the scale) when using the original scoring method.

### Exploratory and confirmatory factor analyses

Results from the CFA for the one-factor structure showed satisfactory fit indices (Table 1), and all factor loadings were greater than 0.30, except for item 1, which was slightly lower (Fig 1). CFA testing a two-factor structure only partially fitted (Table 1), showing four items loading below 0.30 and one above 1 (Heywood case) (Fig 2). Further, the AICs and BIC were lower for the one-factor structure (Table 1), indicating that the one-factor structure fits better than the two-factor structure.

The results of the two-factor EFA (Fig 3) were far from the original two-factor structure (with items 2, 5, 6, 7, and 8 belonging to one factor; items 3, and 4 to another; and item 1 loading below 0.30).

### Reliability

The questionnaire showed a good internal consistency, with alpha values above 0.7 in both the non-clinical and clinical sample. Test-retest reliability evaluation with a subsample of the non-clinical sample showed moderate reliability for the RFQ-8 (Table 2).

Only removing the item 1 seems to improve internal consistency, but it does it in such a way ( $\alpha$  improved to 0.779 in non-clinical sample, and to 0.847 in clinical sample) that the decision was made to test convergent and know-groups validity maintaining all the items, without removing low loading items, and those that overlap in content, as it was made in studies using the RFQ-6 [28, 29]

### Convergent validity

**Mentalizing-related constructs.** In the non-clinical sample, RFQ-8 correlated moderately positively with the IPO-83 Identity Diffusion scale and with the Total TAS-20 scores. On the other hand, it showed a weak negative correlation with the MAAS and with the IRI Perspective Taking scale (Table 3).

In the clinical sample, the RFQ-8 showed a moderate to strong correlation in the expected direction with all the hypothetical mentalizing-related constructs (Table 3).

Correlation with psychopathology measures.

The RF-8 showed significant correlations with severity of general psychopathology, as measured by the Global Severity Index of the SCL-90, both in the clinical and in the non-clinical sample. It also correlated with each of the specific symptom scale in the clinical sample, and with many of them (except somatization, depression, hostility and phobic anxiety; but with

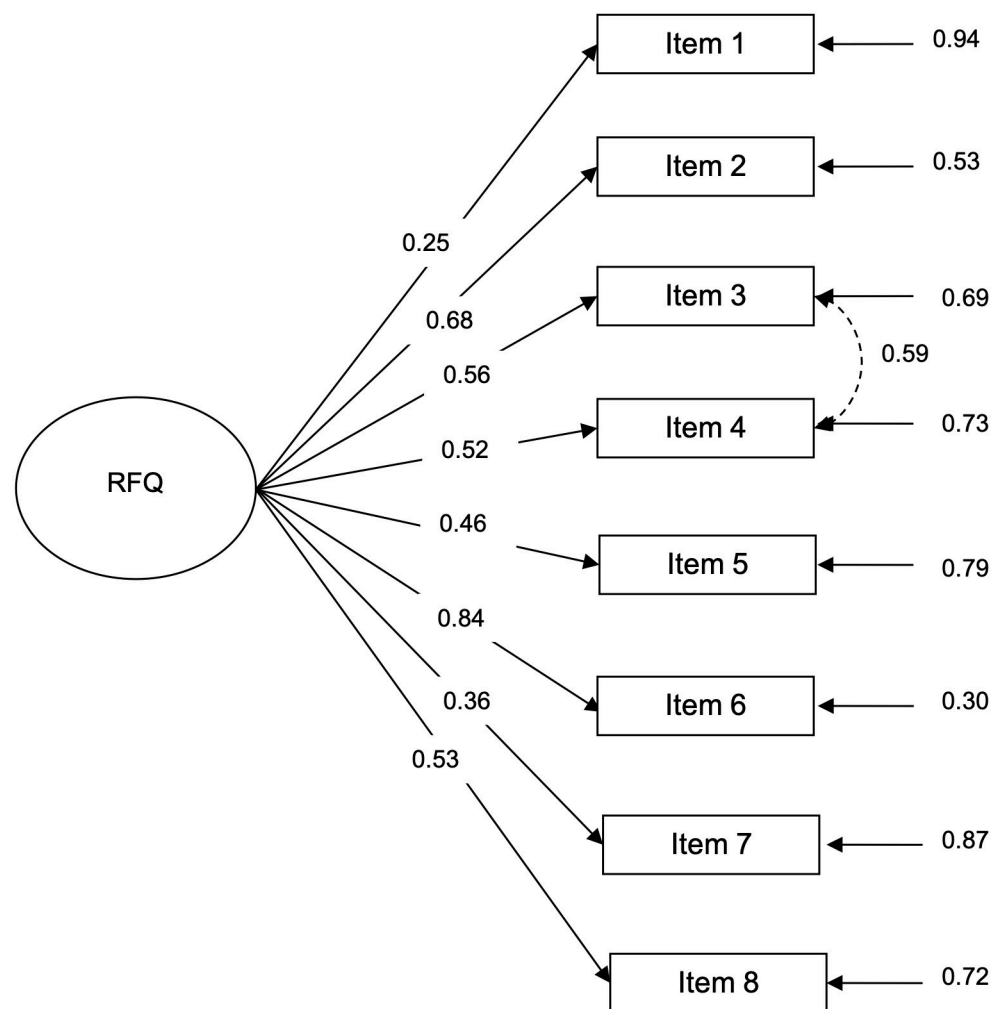
**Table 1. CFA in non-clinical sample using a one-factor and a two-factor model: goodness-of-fit indices and comparison criterions.**

	N	$\chi^2$	df	$\chi^2/df$	RMSEA (90%CI)	TLI	CFI	AIC	BIC	ABIC
RFQ 1f	602	53.46	19	2.81	0.055 (0.038–0.073)	0.96	0.97	17954.02	18064.02	17984.65
RFQ 2f	602	45.73	14	3.27	0.061 (0.042–0.082)	0.95	0.97	17956.28	18088.28	17993.05

Note: RFQ 1f: One-factor model; RFQ 2f: two-factor model;  $\chi^2$ : Chi Square; df: Degrees of freedom; RMSEA: Root Mean Square Error of Approximation; CI: Confidence Interval; TLI: Tucker-Lewis Index; CFI: Comparative Fit Index; AIC: Akaike Information Criterion; BCI: Bayesian Information Criterion; ABIC: Adjusted BIC.

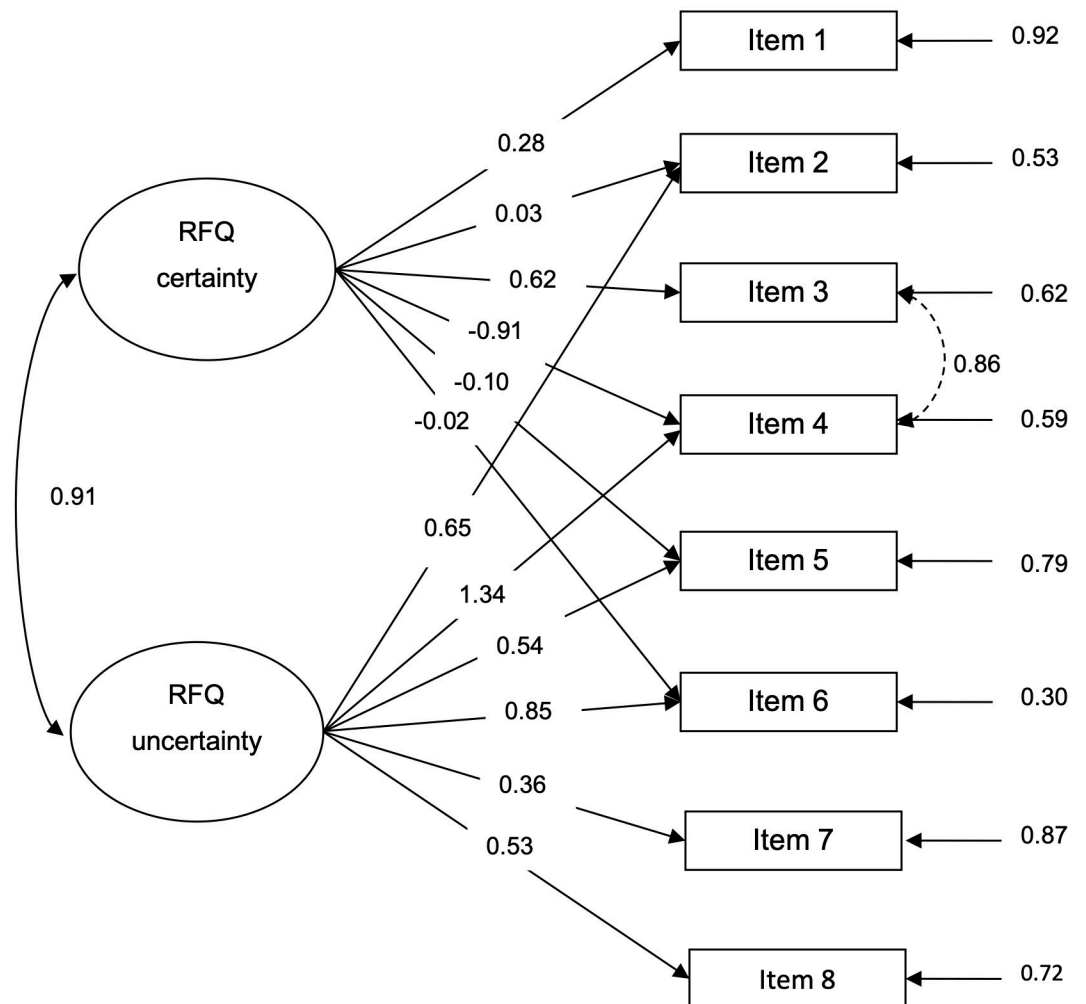
<https://doi.org/10.1371/journal.pone.0274378.t001>

absolute values of the correlation close to 0.4) in the non-clinical sample. On the other hand, the RFQ-8 showed a moderate correlation with the severity of depression, measured by the BDI-II, but only in the clinical sample (Table 4).



**Fig 1. Confirmatory factor analysis for the one-factor structure in the non-clinical sample.** Note: The standardized factor loadings and error variances are shown. As in Spitzer et al. [28], Müller et al. [29], and Wozniak-Prus et al. [38] studies, the error of items 3 and 4 were allowed to covariate.

<https://doi.org/10.1371/journal.pone.0274378.g001>



**Fig 2. Confirmatory factor analysis for the two-factor structure in the non-clinical sample.** Note: The standardized factor loadings, error variances and covariance among exogenous variables are shown. As in Spitzer et al. [28], and Müller et al [29] studies, the error of items 3 and 4 were allowed to covariate.

<https://doi.org/10.1371/journal.pone.0274378.g002>

The RFQ-8 correlated with the severity of personality dysfunction in the non-clinical sample. Finally, the RF-8 correlated with general interpersonal difficulties in the clinical sample (Table 4).

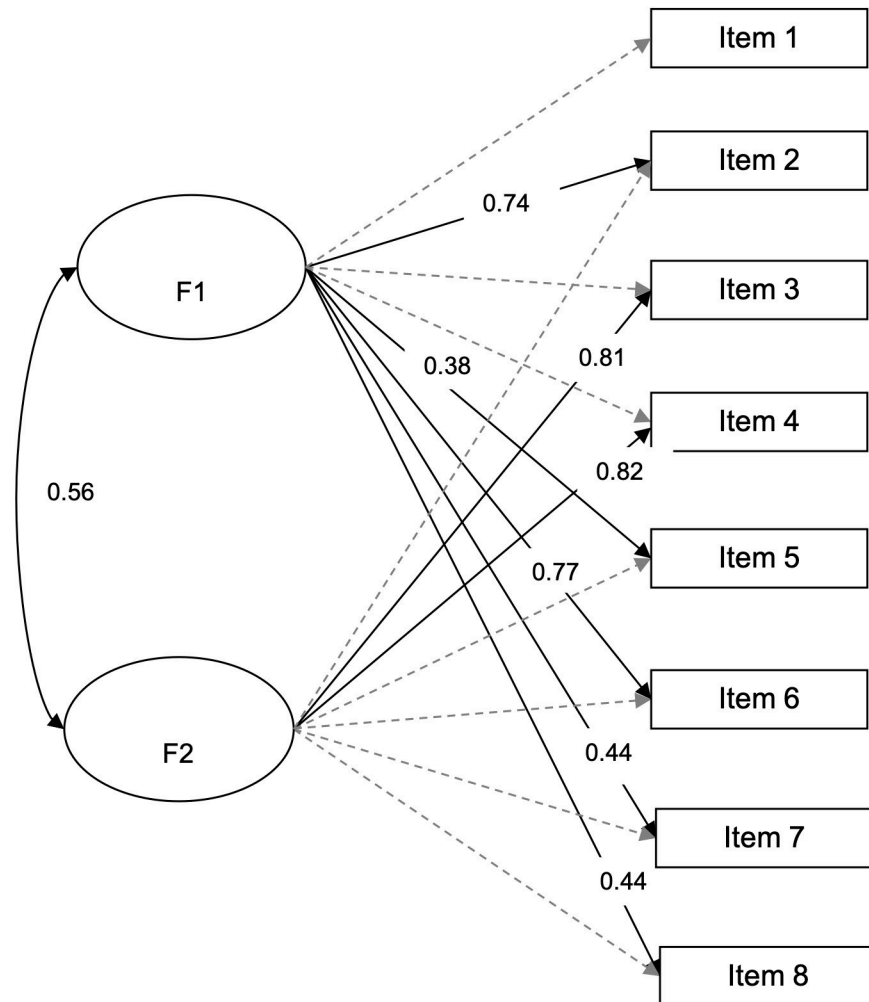
Group differences.

RFQ-8 mean values also discriminated between the non-clinical group and the clinical group of individuals with personality disorder (Table 5).

## Discussion and conclusions

### Translation and equivalence

According to the thorough translation procedure, the resulting version of the questionnaire apparently shows an adequate conceptual, semantic and operational equivalence with the original questionnaire. However, the new version inherits from the original some issues related to face and content validity, previously pointed out by Müller et al. [29]. The RFQ-8 includes only one question assessing thinking about other people; and seven questions regarding



**Fig 3. Exploratory factor analysis for the two-factor structure in the non-clinical sample using Promax rotation.** Note: The factor loadings are shown, and those smaller than 0.30 are grayed out. The percentage of variance explained by the two factors was 53.01%.

<https://doi.org/10.1371/journal.pone.0274378.g003>

oneself, including five about feelings and two about thinking about behaviors. Items 2 and 6 are equivalent in content [27]. The RFQ-8 shows a loss of content validity with respect to its previous version, the RFQ-54.

Additionally, all items (except item 7) make statements about a state of uncertainty, resulting in a loss of face validity of the RFQc scale when applying the original scoring method, and probably, in a lack of capacity of the RFQ-8 to measure a hypermentalization pole when applying the new scoring method [29]. Maybe the RFQ-8 could have more adequately represented

**Table 2. Reliability indices (internal consistency and temporal stability) for non-clinical and clinical groups.**

	Cronbach's Alpha	Test-Retest ICC
Non-clinical sample	0.763	0.746*
Clinical sample	0.783	-

\*p<0.001

<https://doi.org/10.1371/journal.pone.0274378.t002>

**Table 3. Correlations between RFQ and measures of mentalizing-related constructs among non-clinical and clinical sample.**

		IPO		MAAS+	PR_IRI+	TAS		
		ID+	PD			T+	DIF	DDF
<b>RFQ</b>								
<b>Non-clinical</b>								
	n	258		254	323	323		
	rho	0.581*	0.496*	-0.286*	-0.223*	0.487*	0.565*	0.296*
<b>Clinical</b>								
	n	41		41	41	41		
	rho	0.817*	0.692*	-0.448**	-0.424**	0.685*	0.770*	0.413***

Note 1: IPO: Inventory of Personality Organization; ID: Identity Diffusion Scale; PD: Primitive Defenses Scale; MAAS: Mindful Attention Awareness Scale; PT\_IRI: Perspective Taking Scale of Interpersonal Reactivity Index; TAS: Toronto Alexithymia Scale; T: Total; DIF: Difficulty Identifying Feelings Subscale; DDF: Difficulty Describing Feelings Subscale; +: Hypothetical Constructs to test; rho: Spearman Correlation Coefficient.

Note 2:

\*  $p < 0.001$ ;

\*\*  $p < 0.01$

<https://doi.org/10.1371/journal.pone.0274378.t003>

the complexity of the RF definition if its items had not been selected from the RFQ-54 only by purely mathematical criteria.

### Characteristics of the samples

The non-clinical sample has a large sample size [111], providing sufficient robustness for the factor analyses, and other statistical analyses; and the clinical sample size is suitable for the analysis of convergent and known-groups validity. The overrepresentation of the female population in both samples might cause problems in the application of the questionnaires to a more homogeneous population. However, there is no previous evidence of significant differences in RF between both sexes [36]. In the non-clinical sample, a young population, with a medium-high educational level, is overrepresented, which would require an evaluation of the performance of the scales when applying the RFQ-8 to a more diverse population. The clinical sample differs significantly from the non-clinical sample in terms of age (higher in the former), and educational level (more diverse in the former). These differences should be considered when interpreting known-groups validity.

### Floor and ceiling effects

The absence of ceiling and floor effects when using the new scoring method is not surprising, and may be indicative of the capacity of the questionnaire to discriminate a wider range of mentalization difficulties (i.e., hypomentalization). Furthermore, it makes sense that the analysis revealed a slight floor effect for the RFQc scale, especially when applying the RFQ-8 in the nonclinical population, since this scale is expected to better capture severe mentalization deficits [27]. Original rescoring system was designed to capture extremes.

### Factorial structure of the questionnaire

Replicating the results of more recent studies [28, 29, 38], our study provides consistent evidence that a one-factor model adequately explains the observed covariation of RFQ-8 responses when using the new scoring system (proposed to evaluate concerns about the double-scoring procedure, the factor structure of the RFQ, and the validity of the RFQc scale [29]), challenging the capacity of the RFQ-8 to measure two different mentalization failures. In



**Table 4. Correlations between RFQ-8 and measures of psychopathology among non-clinical and clinical sample.**

		RFQ			
		Non-clinical		Clinical	
		n	rho	n	Rho
<b>SCL-90-R</b>		293		40	
	Somatization		0.391*		0.457***
	Obsessive-Compulsive		0.416*		0.590*
	Interpersonal Sensitivity		0.409*		0.512**
	Depression		0.390*		0.485***
	Anxiety		0.423*		0.542**
	Anger-Hostility		0.361*		0.578*
	Phobic Anxiety		0.356*		0.542**
	Paranoid Thought		0.406*		0.585*
	Psychoticism		0.377*		0.564**
	Global Severity Index+		0.442*		0.548**
	Positive Symptoms Total		0.424*		0.491***
<b>BDI-II+</b>		300	0.330*	40	0.408**
<b>PID-5-BF</b>		311			
	PID5 Overall+		0.477*		
	Negative Affect		0.484*		
	Detachment		0.235*		
	Antagonism		0.253*		
	Disinhibition		0.416*		
	Psychoticism		0.354*		
<b>IIP-32</b>		293		40	
	IIP Overall+		0.382*		0.450***

Note 1: SCL-90-R: Symptom Checklist 90 Revised; BDI-II: Beck Depression Inventory II; PID-5-BF: Personality Inventory for DSM 5 Brief Form; IIP-32: Inventory of Interpersonal Problems 32. +: Hypothesized correlations; rho: Spearman Correlation Coefficients.

Note 2:

\* $p < 0.0001$ ;

\*\* $p < 0.001$ ;

\*\*\* $p < 0.01$

<https://doi.org/10.1371/journal.pone.0274378.t004>

contrast with other validation studies [4, 27, 35–37] using the original scoring method, goodness-of-fit indices suggest that the RFQ-8 fails to adjust to a two-factor model when using CFA applying both the new and the original scoring methods (S1 Fig).

In our study, when one-factor CFA is performed using the new scoring system, item 1 is the only one that exhibits a factor loading below 0.30 (0.25), reproducing one study evaluating

**Table 5. Group differences between the non-clinical and the clinical samples.**

	N	RFQ-8		
		Mean	SD	p value
<b>Non-clinical</b>	602	3.86	1.09	
<b>Personality disorders</b>	41	4.74	1.32	
				<0.0001

<https://doi.org/10.1371/journal.pone.0274378.t005>

unidimensional RFQ [38], but not others [28, 29]. It may be worth noting that this is the only item that addresses mentalization about the others [29].

In Spitzer et al study [28] item 7 showed a negligible factor loading (0.14), and in Müller et al study [29] showed the lowest factor loadings in clinical sample (0.34) and both non-clinical samples (0.45, and 0.49). Besides, the authors proposed that items 3 and 4 largely overlap in their content (so their errors were allowed to covariate in both CFA). In those studies, the decision was made to remove items 4 and 7 [28, 29] to produce the RFQ-6. Although in Müller et al. study [29] all analyses were performed using both the 6-item and the 8-item versions leading to similar results. In our study the factor analyses made by the previous authors were replicated, including the decision to allow error correlations between items 3 and 4, as was suggested by the results of Lagrange multiplier test. It should be considered that the original authors suggested to limit the number of possible error correlations to a minimum, allowing only error correlations between items similar in formulation or meaning, but they explicitly cited items 2 and 6 (without excluding other possibilities) as overlapping in content [27]. In our study, only in the originally proposed two-dimensional CFA using double-scoring, error correlations between items 2 and 6, and between items 3 and 4 were allowed (S1 Fig).

Maintaining a more conservative position than some previous authors [28], we decided to keep all the items of the RFQ-8. Although the internal consistency of the questionnaire improved only if item 1 was removed, it did it in a scarcely relevant way, and the advantages of keeping a unified questionnaire in different languages are clear.

## Reliability

The results of the present study demonstrate that the Spanish version of the RFQ8 is a reliable instrument, with a good internal consistency and acceptable temporal stability. This replicates the findings of previous validation studies using the new scoring method [28, 29, 38].

## Convergent validity

One of the inevitable limitations of the study is that criterion validation is not possible, as the RFS, the gold standard in the measure of mentalization, is not available in Spanish. Besides, due to the lack of availability of other specific mentalizing self-reported questionnaires (like the MZQ, the MentS, the CAMSQ, or the MMQ) in Spanish, convergent validity relies on comparing the RFQ with less specific related constructs, as has been usual in previous validation studies.

The hypothesis of the RFQ-8 as correlating with the degree of identity diffusion both in the nonclinical and in the clinical samples is confirmed. The results are consistent with previous quantitative studies linking hypomentalization with identity diffusion and conditions where identity diffusion predominates [27, 57]. Such studies conceive hypomentalization as a causal or mediating variable. Nevertheless, mentalizing capacity and identity are probably related to each other in several ways. The achievement of mentalizing capacity enables the establishment of a sense of identity [119]. Moreover, mentalizing and identity disturbances are intertwined in personality disordered patients [120]. Finally, exploring identity demands thinking in terms of mental states regarding oneself and, according to the object relations model underlying the IPO-83 [121], regarding the others.

The RFQ-8 is also correlated with the Primitive Defenses scale of the IPO-83. Such correlation was somewhat expected, given the closely related nature of identity diffusion and primitive defenses in the model underpinning the IPO-83, as reflected in its original validation study factor analysis [91]. Some primitive defense mechanisms, such as splitting, involve inaccessibility at one point to a substantial part of internal experience about oneself or the others,

thereby making difficult to understand oneself or the others in terms of elaborate and complex intentional mental states. As expected, the correlations of the RFQ-8 with identity diffusion and primitive defenses were higher in the personality disordered patients sample.

Replicating previous research [4, 35, 37], this study confirms the hypothesis of a correlation between hypomentalization and alexithymia, as reflected by the total alexithymia scores. Alexithymia might be conceived as an indirect measure of lack of emotional mentalization about oneself. The content of the RFQ-8 is biased toward these two dimensions of mentalizing, since 5 out of its 8 items assess the ability to mentalize feelings or emotions about oneself. As expected, the correlations were higher in the clinical sample. Additionally, the RFQ-8 specifically correlated with the Difficulty in Identifying Feelings subscale, the scale that most accurately reflects the reported dimensions of mentalization. The Difficulty in Expressing Feelings subscale of the TAS-20 assesses a general inhibited behavior when expressing feelings, hence its lower correlation with the RFQ-8 is understandable.

In previous studies using Perspective Taking scale [27, 37], only a weak to very weak correlation with hypomentalization (measured by RFQ scale) was found in a non-clinical sample. In a recent study conducted on a non-clinical sample in the US, Müller et al. [29] found unidimensional RFQ-8 correlated poorly with measures of mentalizing others, warning about the possibility that the instrument may not adequately measure mentalization regarding the others. In our study, a moderate negative correlation was found in the clinical sample, but in the non-clinical sample the correlation was weak, although stronger than in the original validation study [27]. These inconsistent correlations result in the RFQ-8's ability to assess other-focused dimension of mentalization remains controversial. Perspective taking denotes a conscious effort to put oneself in the other's position and adopt his or her point of view. In this regard, mentalizing about the other is underrepresented in the RFQ-8, as only item 1 accounts for it, which could explain weaker correlations than initially expected.

The same pattern of correlations reproduces with the measure of mindfulness: a moderate negative correlation with RFQ-8 in clinical sample, and a weak correlation in non-clinical sample. Previous studies using MAAS or other scales measuring mindfulness showed weak correlations of the construct with hypomentalization in non-clinical samples. All these findings could be accounted for by differing emphases between the mentalization and mindfulness approaches. Mindfulness, unlike mentalizing, defines a perceptual and non-evaluative activity, not only encompassing mental states, and strictly focused on the present moment.

The study confirms that the RFQ-8 shows a strong correlation with several measures of psychopathology, either global or specific. The correlation is particularly strong when applying the questionnaire to personality-disordered subjects. In this population there is not only an association between hypomentalization and severity of general psychopathology (measured by the Global Severity Index of the SCL-90-R), and severity of depression (measured by the BDI-II), but also, a correlation is appreciated with each of the specific psychopathology measurement scales of the SCL-90-R, that cover a wide variety of symptomatic dimensions, including depression, anxiety, phobias, obsessive-compulsiveness, somatization, hostility, paranoid ideation, feelings of inferiority and inadequacy, and psychoticism.

In the nonclinical sample, the RFQ-8 correlated adequately with the overall severity of personality dysfunction and with the negative affect and disinhibition domains of the PID-5-BF. This study confirms the correlation between hypomentalization and personality difficulties found in previous studies, which mainly showed moderate to strong correlations with uncertainty scale [4, 27] or unidimensional RFQ [29, 38]. Only one study [35] that used a clinician-rated questionnaire to measure personality disfunction in individuals with borderline personality disorder showed a weak correlation with uncertainty scale.

The correlation of the RFQ-8 with negative affectivity and disinhibition in our study replicates previous studies [29, 38], and it is consistent with the assumption that inadequate mentalizing capacity (i.e., hypomentalization) produces emotional dysregulation and impulsivity [1]. However, some authors have advised that several RFQ-8 items directly assess impulsive behaviors in the context of emotional lability that may be the result of difficulties in mentalization, but may have other causes. This may produce an artificial inflation of correlations between the RFQ and indicators of personality pathology [29]. A communality analysis by Müller et al demonstrated that, although the RFQ-8 reflected impairments in mentalizing, 30% of the observed associations between the RFQ-8 and indicators of personality dysfunction were due to variance shared with measures of emotional lability and impulsivity [29]. Furthermore, an item-level analysis suggested that items 3, 4, 5, and 8 converged with measures of emotional lability and impulsivity rather than with measures of mentalization, and thus those items may be responsible for the artificially inflated correlations [29]. These findings should be considered when interpreting the associations of the RFQ-8, understood as a measure of mentalization, with other constructs, given that the instrument could also be measuring a combination of impulsivity and emotional dysregulation that is partly independent of mentalization difficulties, and that could act as a confounding factor.

The correlation between hypomentalization and the presence of interpersonal problems was moderate in the clinical sample, in agreement with previous findings using the IIP-32 [29]. The correlation was significant but weak when tested in a non-clinical sample for the first time. Some authors have suggested that hypomentization predicts interpersonal problems in individuals with personality disorder, but do it only indirectly, via emotional dysregulation and impulsivity [51].

Consistent with the hypothesis that personality disorders are mainly based on failures of mentalization [122], the group of individuals with personality disorders exhibit significantly higher mean scores of the questionnaire than the non-clinical group.

In the absence of measures in our language that directly measure mentalization, our study, in addition to assessing the fairly well-established relationship between hypomentalization and psychopathology, advances in the sense of comparing the questionnaire with mentalization related constructs, such as those used in the initial validation studies [4, 27, 32]. These constructs are now evaluated using the RFQ-8 as a one-factor questionnaire using the new scoring method. Some of them showed adequate correlations in the clinical sample with measures of cognitive empathy, besides mentalization about the others underrepresentation in the RFQ-8; and mindfulness, besides mentioned differing emphases in the approaches. Furthermore, we have understood alexithymia, and specially its DIF scale, as one of this closed and more genuine mentalization related constructs, measuring lack of emotional mentalizing about oneself (and not only as a measure of psychopathology). In a similar way, normal identity exploration is not possible without the capacity of genuinely mentalizing about oneself and the others. It is therefore not surprising that both constructs correlate well with the RFQ in both clinical and non-clinical samples and must be considered beyond their role as psychopathological indicators.

The generalized presence of significantly stronger correlations in the clinical sample than in the non-clinical sample for all the constructs under study suggests a higher reliability and validity of the questionnaire in the group of patients with personality disorders, and points to the need for validation studies with larger samples for this group.

**RFQ-8 as an off-line measure.** The RFQ8 is an offline measure of mentalization. It requests the subject to reflexively assess his or her own general mentalizing ability, without ensuring a current interpersonal, affectively relevant context for the assessment, and without assuring a minimum level of stress to promote the emergence of mentalization failures. The

RFQ-8 has a limited capacity to include the dynamic and contextual elements of mentalization [7]. It provides us with a global picture of the subject's mentalizing ability focused on one of its deficits, hypomentalization. The use of a standardized stress-inducing tasks [18] enables the assessment of automatic mentalization with other instruments [19], but probably this is not the case with the RFQ-8. Nevertheless, it is an easy-to-use instrument that provides a direct measure of hypomentalization. Yet, the need to validate measures in Spanish of the other mentalization deficit, hypermentalization, remains. In response to growing evidence, original developers of RFQ are in the process of validating a new version of the RFQ with a hypermentalizing scale to replace the certainty scale. We are awaiting their and others findings with this measure.

## Conclusions

In summary, the Spanish version of the RFQ-8 is an off-line self-report reliable instrument with an adequate construct validity. The present study suggest a one-factor structure of the questionnaire. Using the new scoring method proposed, the RFQ-8, as an unidimensional questionnaire aimed at measure hypomentalization, correlates with several measures of psychopathology in clinical and non-clinical population; and with diverse mentalization related constructs, with stronger correlations in clinical sample. Design of a new scale to measure hypermentalization, with adequate face validity, and free of double-scoring problems is encouraged. Further research is needed in our community with measures of hypomentalization and hypermentalization with larger samples of individuals with personality disorders, and another psychopathology.

## Supporting information

**S1 Appendix. The reflective functioning questionnaire.**

(PDF)

**S2 Appendix. Spanish version of the RFQ-8.**

(PDF)

**S1 Fig. Originally proposed two-dimensional CFA model using double-scoring in non-clinical sample.**

(PDF)

**S1 Table. Polychoric correlations among items using the original scoring method.**

(PDF)

**S2 Table. Reliability indices (internal consistency and temporal stability), using double-scoring, for non-clinical and clinical groups.**

(PDF)

**S3 Table. Group differences between the non-clinical and the clinical samples, using double-scoring.**

(PDF)

**S4 Table. Correlations between RFQc and RFQu and measures of mentalizing-related constructs among non-clinical and clinical sample.**

(PDF)

**S5 Table. Correlations between RFQc and RFQu and measures of psychopathology among non-clinical and clinical sample.**

(PDF)



**S1 Dataset. Data used in the study.**  
(XLSX)

## Author Contributions

**Conceptualization:** Eduardo Ruiz-Parra.

**Formal analysis:** Amaia Bilbao.

**Funding acquisition:** Miguel Ángel González-Torres.

**Investigation:** Eduardo Ruiz-Parra, Guadalupe Manzano-García, Roberto Mediavilla, Beatriz Rodríguez-Vega, Guillermo Lahera, Ana I. Moreno-Pérez, Alberto M. Torres-Cantero, Juan Rodado-Martínez.

**Methodology:** Eduardo Ruiz-Parra.

**Project administration:** Miguel Ángel González-Torres.

**Software:** Eduardo Ruiz-Parra.

**Supervision:** Eduardo Ruiz-Parra, Miguel Ángel González-Torres.

**Writing – original draft:** Eduardo Ruiz-Parra.

**Writing – review & editing:** Eduardo Ruiz-Parra, Miguel Ángel González-Torres.

## References

1. Fonagy P, Gergely G, Jurist EL, Target M. Affect regulation, mentalization and the development of the self. New York: Other Press; 2002.
2. Bateman A, Fonagy P. Mentalization-based treatment for borderline personality disorder. A practical guide. Oxford (GB): Oxford University Press; 2006.
3. Sharp C, Pane H, Ha C, Venta A, Patel AB, Sturek J, et al. Theory of mind and emotion regulation difficulties in adolescents with borderline traits. *J Am Acad Child Adolesc Psychiatry*. 2011 Jun; 50(6):563–73.e1. <https://doi.org/10.1016/j.jaac.2011.01.017> PMID: 21621140
4. Badoud D, Luyten P, Fonseca-Pedrero E, Eliez S, Fonagy P, Debbané M. The French version of the Reflective Functioning Questionnaire: validity data for adolescents and adults and its association with non-suicidal self-injury. *PLOS One*. 2015 Dec 29; 10(12):e0145892. <https://doi.org/10.1371/journal.pone.0145892> PMID: 26714319
5. Bateman A, Fonagy P. Randomized controlled trial of outpatient mentalization-based treatment versus structured clinical management for borderline personality disorder. *Am J Psychiatry*. 2009 Dec; 166(12):1355–64. <https://doi.org/10.1176/appi.ajp.2009.09040539> PMID: 19833787
6. Fonagy P, Steele M, Steele H, Moran GS, Higgitt AC. The capacity for understanding mental states: the reflective self in parent and child and its significance for security of attachment. *Infant Ment Health J*. 1991; 12(3):201–18. [https://doi.org/10.1002/1097-0355\(199123\)12:3<201::AID-IMHJ2280120307>3.0.CO;2-7](https://doi.org/10.1002/1097-0355(199123)12:3<201::AID-IMHJ2280120307>3.0.CO;2-7).
7. Luyten P, Malcorps S, Fonagy P, Ensink K. Assessment of mentalization. In: Bateman A, Fonagy P, editors. *Handbook of mentalizing in mental health practice*. 2nd ed. Washington, DC: American Psychiatric Association Publishing; 2019. p. 37–62.
8. Sánchez Pascual M, León Quismondo L, García López de Arenosa A, Mass Hesse J, Fernández Liria A. La medida de la capacidad reflexiva: instrumentos disponibles en castellano y tareas pendientes. [Measuring reflecting capacity: available instruments in Castilian and pending tasks]. *Rev Asoc Esp Neuropsiquiatría*. 2015; 35(127):487–510. <https://doi.org/10.4321/S0211-57352015000300004>
9. Fonagy P, Target M, Steele H, Steele M. *Reflective Functioning Manual, version 5.0, for application to Adult Attachment Interviews*. London: University College London; 1998.
10. George C, Kaplan N, Main M. *The Adult Attachment Interview*. Berkeley: University of California Berkeley; 1985.

11. Slade A, Bernbach E, Grienenberger J, Levy D, Locker A. Addendum to Fonagy, Target, Steele, & Steele Reflective Functioning Scoring Manual for use with the Parent Development Interview. Version 2.0. New York: The City College and Graduate Center of the City University of New York; 2004.
12. Slade A, Aber JL, Bresgi I, Berger B, Kaplan M. The Parent Development Interview—Revised. New York: The City University of New York; 2004.
13. Grienenberger JF, Kelly K, Slade A. Maternal reflective functioning, mother-infant affective communication, and infant attachment: exploring the link between mental states and observed caregiving behavior in the intergenerational transmission of attachment. *Attach Hum Dev*. 2005 Sep; 7(3):299–311. <https://doi.org/10.1080/14616730500245963> PMID: 16210241
14. Zeanah CH, Benoit D, Barton ML, Hirshberg L. Working Model of the Child Interview Coding Manual [unpublished manuscript]. New Orleans: Louisiana State University School of Medicine; 1996.
15. Ensink K, Normandin L, Target M, Fonagy P, Sabourin S, Berthelot N. Mentalization in children and mothers in the context of trauma: an initial study of the validity of the Child Reflective Functioning Scale. *Br J Dev Psychol*. 2015 Jun; 33(2):203–17. <https://doi.org/10.1111/bjdp.12074> PMID: 25483125
16. Target M, Fonagy P, Shmueli-Goetz Y, Schneider T, Datta A. Child Attachment Interview (CAI): Coding and Classification Manual. Version III. London: University College London; 2000.
17. Fertuck EA, Mergenthaler E, Target M, Levy KN, Clarkin JF. Development and criterion validity of a computerized text analysis measure of reflective functioning. *Psychother Res*. 2012; 22(3):298–305. <https://doi.org/10.1080/10503307.2011.650654> PMID: 22320840
18. Kirschbaum C, Pirke KM, Hellhammer DH. The 'Trier Social Stress Test'—a tool for investigating psychobiological stress responses in a laboratory setting. *Neuropsychobiology*. 1993; 28(1–2):76–81. <https://doi.org/10.1159/000119004> PMID: 8255414
19. Finn MTM, Smith CL, McKernan LC, Nash MR. Moving and reflective functioning under stress. *Psychodyn Psychiatry*. 2019 Summer; 47(2):197–214. <https://doi.org/10.1521/pdps.2019.47.2.197> PMID: 31107164
20. Taubner S, Hörz S, Fischer-Kern M, Doering S, Buchheim A, Zimmermann J. Internal structure of the Reflective Functioning Scale. *Psychol Assess*. 2013 Mar; 25(1):127–135. <https://doi.org/10.1037/a0029138> PMID: 22775410
21. Vrouva I, Fonagy P. Development of the Mentalizing Stories for Adolescents (MSA). *J Am Psychoanal Assoc*. 2009; 57(5):1174–9. <https://doi.org/10.1177/0003065109347904>
22. Luyten P, Mayes LC, Nijssens L, Fonagy P. The Parental Reflective Functioning Questionnaire: development and preliminary validation. *PLOS One*. 2017; 12(5):e0176218. <https://doi.org/10.1371/journal.pone.0176218> PMID: 28472162
23. Hausberg MC, Schulz H, Piegler T, Happach CG, Klöpfer M, Brütt AL, et al. Is a self-rated instrument appropriate to assess mentalization in patients with mental disorders? Development and first validation of the Mentalization Questionnaire (MZQ). *Psychother Res J Soc Psychother Res*. 2012; 22(6):699–709. <https://doi.org/10.1080/10503307.2012.709325> PMID: 22867004
24. Dimitrijević A, Hanak N, Altaras Dimitrijević A, Jolić Marjanović Z. The Mentalization Scale (MentS): a self-report measure for the assessment of mentalizing capacity. *J Pers Assess*. 2018 Jun; 100(3):268–80. <https://doi.org/10.1080/00223891.2017.1310730> PMID: 28436689
25. Müller S, Wendt LP, Zimmermann J. Development and Validation of the Certainty About Mental States Questionnaire (CAMSQ): A Self-Report Measure of Mentalizing Oneself and Others. *Assessment*. 2021 Dec 14;10731911211061280. <https://doi.org/10.1177/10731911211061280> PMID: 34905983
26. Gori A, Arcioni A, Topino E, Craparo G, Lauro Grotto R. Development of a new measure for assessing mentalizing: the Multidimensional Mentalizing Questionnaire (MMQ). *J Pers Med*. 2021 Apr; 11(4):305. <https://doi.org/10.3390/jpm11040305> PMID: 33921083
27. Fonagy P, Luyten P, Moulton-Perkins A, Lee Y-W, Warren F, Howard S, et al. Development and validation of a self-report measure of mentalizing: the Reflective Functioning Questionnaire. *PLOS One*. 2016; 11(7):e0158678. <https://doi.org/10.1371/journal.pone.0158678> PMID: 27392018
28. Spitzer C, Zimmermann J, Brähler E, Euler S, Wendt L, Müller S. Die deutsche Version des Reflective Functioning Questionnaire (RFQ): Eine teststatistische Überprüfung in der Allgemeinbevölkerung. [The German version of the Reflective Functioning Questionnaire (RFQ): a psychometric evaluation in the general population]. *Psychother Psychosom Med Psychol*. 2021; 71(3/4):124–31. <https://doi.org/10.1055/a-1234-6317> PMID: 33063306
29. Müller S, Wendt LP, Spitzer C, Masuhr O, Back SN, Zimmermann J. A Critical Evaluation of the Reflective Functioning Questionnaire (RFQ). *J Pers Assess*. 2021. <https://doi.org/10.1080/00223891.2021.1981346> PMID: 34597256
30. Perkins A. Feelings, faces and food: Mentalization in borderline personality disorder and eating disorders [Internet] [DClinPsych]. University of Surrey; 2019. Available from: <https://openresearch.surrey>.

- [ac.uk/esploro/outputs/doctoral/Feelings-Faces-and-Food-Mentalization-in/99513178502346?institution=44SUR\\_INST](https://doi.org/10.1371/journal.pone.0259030).
31. Rogoff S, Moulton-Perkins A, Warren F, Nolte T, Fonagy P. 'Rich' and 'poor' in mentalizing: Do expert mentalizers exist?. *PloS One*. 2021; 16(10):e0259030. <https://doi.org/10.1371/journal.pone.0259030> PMID: 34695171
  32. Sharp C, Steinberg L, McLaren V, Weir S, Ha C, Fonagy P. Refinement of the Reflective Function Questionnaire for Youth (RFQY) Scale B Using Item Response Theory. *Assessment*. 2021;10731911211003971. <https://doi.org/10.1177/10731911211003971> PMID: 33794672
  33. Moulton-Perkins A, Rogoff S, Fonagy P, Luyten P. Development and validation of a new self-report measure of mentalization: The 54-item Reflective Function Questionnaire. Paper presented at: The 17th British Association of Clinical Psychologists/Society for Psychotherapy Research Conference; 2011 May 6-7; Liverpool, UK.
  34. University College London. The Reflective Functioning Questionnaire (RFQ) [Internet]. *Psychoanalysis Unit*. 2018 [cited 2022 May 1]. Available from: <https://www.ucl.ac.uk/psychoanalysis/research/reflective-functioning-questionnaire-rfq>.
  35. Morandotti N, Brondino N, Merelli A, Boldrini A, Vidovich GZD, Ricciardo S, et al. The Italian version of the Reflective Functioning Questionnaire: validity data for adults and its association with severity of borderline personality disorder. *PLOS One*. 2018 Nov 1; 13(11):e0206433. <https://doi.org/10.1371/journal.pone.0206433> PMID: 30383803
  36. Griva F, Pomini V, Gournellis R, Doumos G, Thomakos P, Vaslamatzis G. Psychometric properties and factor structure of the Greek version of Reflective Functioning Questionnaire. *Psychiatriki*. 2020 Jul-Sep; 31(3):216–24. <https://doi.org/10.22365/jpsych.2020.313.216> PMID: 33099462
  37. Mousavi PSS, Vahidi E, Ghanbari S, Khoshroo S, Sakkaki SZ. Reflective Functioning Questionnaire (RFQ): Psychometric properties of the Persian translation and exploration of its mediating role in the relationship between attachment to parents and internalizing and externalizing problems in adolescents. *J Infant Child Adolesc Psychother*. 2021; 20(3):313–30. <https://doi.org/10.1080/15289168.2021.1945721>
  38. Wozniak-Prus M, Gambin M, Cudo A, Sharp C. Investigation of the factor structure of the Reflective Functioning Questionnaire (RFQ-8): One or two dimensions?. *J Pers Assess*. 2022:1–11. <https://doi.org/10.1080/00223891.2021.2014505> PMID: 35015610
  39. Agostini A, Scaiola E, Belluzzi A, Campieri M. Attachment and mentalizing abilities in patients with inflammatory bowel disease. *Gastroenterol Res Pract*. 2019 Dec 5; 2019:7847123. <https://doi.org/10.1155/2019/7847123> PMID: 31885546
  40. Ciccarelli M, Nigro G, D'Olimpio F, Griffiths MD, Cosenza M. Mentalizing failures, emotional dysregulation, and cognitive distortions among adolescent problem gamblers. *J Gambli Stud*. 2021 Mar; 37(1):283–98. <https://doi.org/10.1007/s10899-020-09967-w> PMID: 32720217
  41. Cucchi A, Hampton JA, Moulton-Perkins A. Using the validated Reflective Functioning Questionnaire to investigate mentalizing in individuals presenting with eating disorders with and without self-harm. *PeerJ*. 2018 Oct 29; 6:e5756. <https://doi.org/10.7717/peerj.5756> PMID: 30397541
  42. Gagliardini G, Gullo S, Tinozzi V, Baiano M, Balestrieri M, Todisco P, et al. Mentalizing subtypes in eating disorders: a latent profile analysis. *Front Psychol* 2020 Nov 30; 11:564291. <https://doi.org/10.3389/fpsyg.2020.564291> PMID: 33329192
  43. Macfie J, Zvara BJ, Stuart GL, Kurdziel-Adams G, Kors SB, Fortner KB, et al. Pregnant women's history of childhood maltreatment and current opioid use: the mediating role of reflective functioning. *Addict Behav*. 2020 Mar; 102:106134. <https://doi.org/10.1016/j.addbeh.2019.106134> PMID: 31863966
  44. Sacchetti S, Robinson P, Bogaardt A, Clare A, Ouellet-Courtois C, Luyten P, et al. Reduced mentalizing in patients with bulimia nervosa and features of borderline personality disorder: a case-control study. *BMC Psychiatry*. 2019 May 6; 19(1):134. <https://doi.org/10.1186/s12888-019-2112-9> PMID: 31060534
  45. Perroud N, Badoud D, Weibel S, Nicastrò R, Hasler R, Küng A-L, et al. Mentalization in adults with attention deficit hyperactivity disorder: comparison with controls and patients with borderline personality disorder. *Psychiatry Res*. 2017 Oct; 256:334–341. <https://doi.org/10.1016/j.psychres.2017.06.087> PMID: 28675858
  46. Mariani R, Musetti A, Di Monte C, Danskin K, Franceschini C, Christian C. Maladaptive daydreaming in relation to linguistic features and attachment style. *Int J Environ Res Public Health*. 2021; 19(1):386. <https://doi.org/10.3390/ijerph19010386> PMID: 35010644
  47. De Coninck D, Matthijs K, Luyten P. Depression in Belgian first-year university students: A longitudinal study of self-definition, interpersonal relatedness, mentalizing, and integration. *J Clin Psychol*. 2021 Jul; 77(7):1715–31. <https://doi.org/10.1002/jclp.23149> PMID: 33880755

48. Badoud D, Prada P, Nicastró R, Germond C, Luyten P, Perroud N, et al. Attachment and reflective functioning in women with borderline personality disorder. *J Personal Disord*. 2018 Feb; 32(1):17–30. [https://doi.org/10.1521/pedi\\_2017\\_31\\_283](https://doi.org/10.1521/pedi_2017_31_283) PMID: 28263091
49. Brugnera A, Zarbo C, Compare A, Talia A, Tasca GA, Jong K de, et al. Self-reported reflective functioning mediates the association between attachment insecurity and well-being among psychotherapists. 2021 Feb; 31(2):247–57. <https://doi.org/10.1080/10503307.2020.1762946> PMID: 32429777
50. Cosenza M, Ciccarelli M, Nigro G. The steamy mirror of adolescent gamblers: mentalization, impulsivity, and time horizon. *Addict Behav*. 2019 Feb; 89:156–62. <https://doi.org/10.1016/j.addbeh.2018.10.002> PMID: 30316141
51. Euler S, Nolte T, Constantinou M, Griem J, Montague PR, Fonagy P, et al. Interpersonal problems in borderline personality disorder: associations with mentalizing, emotion regulation, and impulsiveness. *J Personal Disord*. 2021 Apr; 35(2):177–93. [https://doi.org/10.1521/pedi\\_2019\\_33\\_427](https://doi.org/10.1521/pedi_2019_33_427) PMID: 30920937
52. Gambin M, Woźniak-Prus M, Konecka A, Sharp C. Relations between attachment to mother and father, mentalizing abilities and emotion regulation in adolescents. *Eur J Dev Psychol*. 2020; 18(1):18–37. <https://doi.org/10.1080/17405629.2020.1736030>
53. Huang YL, Fonagy P, Feigenbaum J, Montague PR, Nolte T, and London Personality and Mood Disorder Research Consortium. Multidirectional pathways between attachment, mentalizing, and posttraumatic stress symptomatology in the context of childhood trauma. *Psychopathology*. 2020; 53(1):48–58. <https://doi.org/10.1159/000506406> PMID: 32294649
54. Kristiansen VR, Handeland TB, Lau B, Söderström K, Håkansson U, Øie MG. Trauma in childhood and adolescence and impaired executive functions are associated with uncertain reflective functioning in mothers with substance use disorder. *Addict Behav Rep*. 2019 Dec 30; 11:100245. <https://doi.org/10.1016/j.abrep.2019.100245> PMID: 32467834
55. Li ET, Carracher E, Bird T. Linking childhood emotional abuse and adult depressive symptoms: the role of mentalizing incapacity. *Child Abuse Negl*. 2020 Jan; 99:104253. <https://doi.org/10.1016/j.chiabu.2019.104253> PMID: 31812024
56. Morosan L, Fonseca-Pedrero E, Debbané M. Network analysis of reflective functioning and conduct problems during adolescence. *Psychol Violence*. 2020; 10(3):300–11. <https://doi.org/10.1037/vio0000258>
57. Penner F, Gambin M, Sharp C. Childhood maltreatment and identity diffusion among inpatient adolescents: the role of reflective function. *J Adolesc*. 2019 Oct; 76:65–74. <https://doi.org/10.1016/j.adolescence.2019.08.002> PMID: 31472427
58. Salaminios G, Morosan L, Toffel E, Tanzer M, Eliez S, Badoud D, et al. Associations between schizotypal personality features, mentalizing difficulties and thought problems in a sample of community adolescents. *Early Interv Psychiatry*. 2021 Jun; 15(3):705–15. <https://doi.org/10.1111/eip.13011> PMID: 32573985
59. Scandurra C, Dolce P, Vitelli R, Esposito G, Testa RJ, Balsam KF, et al. Mentalizing stigma: reflective functioning as a protective factor against depression and anxiety in transgender and gender-nonconforming people. *J Clin Psychol*. 2020 Sep; 76(9):1613–1630. <https://doi.org/10.1002/jclp.22951> PMID: 32222104
60. Berthelot N, Lemieux R, Garon-Bissonnette J, Lacharité C, Muzik M. The protective role of mentalizing: Reflective functioning as a mediator between child maltreatment, psychopathology and parental attitude in expecting parents. *Child Abuse Negl*. 2019 Sep; 95:104065. <https://doi.org/10.1016/j.chiabu.2019.104065> PMID: 31255871
61. Nigro G, Matarazzo O, Ciccarelli M, D'Olimpio F, Cosenza M. To chase or not to chase: A study on the role of mentalization and alcohol consumption in chasing behavior. *J Behav Addict*. 2019; 8(4):743–53. <https://doi.org/10.1556/2006.8.2019.67> PMID: 31891312
62. Manzano-García G, Ayala-Calvo JC, Desrumaux P. Entrepreneurs' Capacity for Mentalizing: Its Influence on Burnout Syndrome. *Int J Environ Res Public Health*. 2020; 18(1). <https://doi.org/10.3390/ijerph18010003> PMID: 33374922
63. Bezerra GP, Laskoski PB, Terra L, Ramos-Lima LF, Serralta FB, Hauck S. The association between reflective function and global functionality in patients with borderline personality disorder. *Australas Psychiatry Bull R Aust N Z Coll Psychiatr*. 2021; 29(6):586–9. <https://doi.org/10.1177/1039856220981801> PMID: 33380157
64. Schwarzer NH, Nolte T, Fonagy P, Gingelmaier S. Mentalizing mediates the association between emotional abuse in childhood and potential for aggression in non-clinical adults. *Child Abuse Negl*. 2021; 115(can, 7801702):105018. <https://doi.org/10.1016/j.chiabu.2021.105018> PMID: 33676103
65. Green J, Berry K, Danquah A, Pratt D. Attachment security and suicide ideation and behaviour: The mediating role of reflective functioning. *Int J Environ Res Public Health*. 2021 Mar 17; 18(6):3090. <https://doi.org/10.3390/ijerph18063090> PMID: 33802833

66. Bordoagni G, Fino E, Agostini A. Burnout, attachment and mentalization in nursing students and nurse professionals. *Healthc Basel Switz*. 2021; 9(11). <https://doi.org/10.3390/healthcare9111576> PMID: 34828621
67. Kahya Y, Munguldar K. Difficulties in emotion regulation mediated the relationship between reflective functioning and borderline personality symptoms among non-clinical adolescents. *Psychol Rep*. 2022;332941211061072. <https://doi.org/10.1177/00332941211061072> PMID: 35048764
68. Carrera S, Pandolfi G, Cappelletti JY, Padoani W, Salcuni S. Oriented mentalization-based treatment for borderline personality disorder patients: preliminary results at Camposampiero Mental Health Center. *Res Psychother Milano*. 2018 Dec 18; 21(3):336. <https://doi.org/10.4081/ripppo.2018.336> PMID: 32913772
69. De Meulemeester C, Vansteelandt K, Luyten P, Lowyck B. Mentalizing as a mechanism of change in the treatment of patients with borderline personality disorder: a parallel process growth modeling approach. *Personal Disord*. 2018 Jan; 9(1):22–9. <https://doi.org/10.1037/per0000256> PMID: 28569524
70. Midgley N, Cirasola A, Austerberry C, Ranzato E, West G, Martin P, et al. Supporting foster carers to meet the needs of looked after children: A feasibility and pilot evaluation of the Reflective Fostering Programme. *Dev Child Welf*. 2019; 1(1):41–60. <https://doi.org/10.1177/2516103218817550>
71. Frolli A, Ricci MC, Tortorelli FA, Cavallaro A, Valenzano L, Rega A, et al. Emotional education in early onset schizophrenia and Asperger's syndrome. *Behav Sci (Basel)*. 2020 Aug 29; 10(9):131. <https://doi.org/10.3390/bs10090131> PMID: 32872431
72. Eskildsen A, Reinholt N, van Bronswijk S, Brund RBK, Christensen AB, Hvenegaard M, et al. Personalized psychotherapy for outpatients with major depression and anxiety disorders: Transdiagnostic versus diagnosis-specific group cognitive behavioural therapy. *Cogn Ther Res*. 2020; 44(5):988–1001. <https://doi.org/10.1007/s10608-020-10116-1>
73. Halstensen K, Gjestad R, Luyten P, Wampold B, Granqvist P, Stalsett G, et al. Depression and mentalizing: A psychodynamic therapy process study. *J Couns Psychol*. 2021; 68(6):705–18. <https://doi.org/10.1037/cou0000544> PMID: 34351181
74. Jewell T, Herle M, Serpell L, Eivors A, Simic M, Fonagy P, et al. Attachment and mentalization as predictors of outcome in family therapy for adolescent anorexia nervosa. *Eur Child Adolesc Psychiatry*. 2021;(9212296). <https://doi.org/10.1007/s00787-021-01930-3> PMID: 34967934
75. Sonntag M, Russell J. The mind-in-mind study: A pilot randomized controlled trial that compared modified mentalization based treatment with supportive clinical management for patients with eating disorders without borderline personality disorder. *Eur Eat Disord Rev J Eat Disord Assoc*. 2022; 30(3):206–20. <https://doi.org/10.1002/erv.2888> PMID: 35132749
76. Handeland TB, Kristiansen VR, Lau B, Håkansson U, Øie MG. High degree of uncertain reflective functioning in mothers with substance use disorder. *Addict Behav Rep*. 2019 May 18; 10:100193. <https://doi.org/10.1016/j.abrep.2019.100193> PMID: 31193539
77. Carlone C, Milan S. Does your child need therapy? Maternal reflective functioning and perceived need for and use of child mental health treatment. *Attach Hum Dev*. 2020 Mar 13;1–18. <https://doi.org/10.1080/14616734.2020.1734641> PMID: 32167022
78. Malda Castillo J, Browne C, Perez-Algorta G. Examining the association between mentalizing and parental mental health in a sample of caregivers of children with asthma. *Couple Fam Psychol Res Pract*. 2020; 9(1):33–44. <https://doi.org/10.1037/cfp0000132>
79. Jovanc Evic A, Tos Ic Radev M, Stefanovic Stanojevic T. Association between parent's and adult offspring's mentalizing capacity: The moderating role of child's temperament. *Scand J Psychol*. 2021; 62(5):699–708. <https://doi.org/10.1111/sjop.12770> PMID: 34462925
80. First MB, Gibbon M, Spitzer RL, Williams JBW, Benjamin LS. *Entrevista Clínica Estructurada para los Trastornos de Personalidad del Eje II del DSM-IV. [Structured Clinical Interview for DSM-IV Axis II Personality Disorders]*. Barcelona: Masson; 1997.
81. Streiner DL, Norman GR, Cairney J. *Health measurement scales. A practical guide to their development and use*. 5th. ed. Oxford (GB): Oxford University Press; 2015.
82. Guillemin F, Bombardier C, Beaton D. Cross-cultural adaptation of health-related quality of life measures: literature review and proposed guidelines. *J Clin Epidemiol*. 1993 Dec; 46(12):1417–32. [https://doi.org/10.1016/0895-4356\(93\)90142-n](https://doi.org/10.1016/0895-4356(93)90142-n) PMID: 8263569
83. Davis MH. A multidimensional approach to individual difference in empathy. *JSAS Catalog Select Doc Psychol*. 1980; 10(85).
84. Pérez-Albéniz A, de Paúl J, Etxeberria J, Montes MP, Torres E. Adaptación de Interpersonal Reactivity Index (IRI) al español. [Spanish adaptation of the Interpersonal Reactivity Index]. *Psicothema*. 2003; 15(2):267–72. Available from: <http://www.psicothema.com/psicothema.asp?id=1056>.



85. Bagby RM, Parker JDA, Taylor GJ. The twenty-item Toronto Alexithymia scale—I. Item selection and cross-validation of the factor structure. *J Psychosom Res.* 1994 Jan; 38(1):23–32. [https://doi.org/10.1016/0022-3999\(94\)90005-1](https://doi.org/10.1016/0022-3999(94)90005-1) PMID: 8126686
86. Bagby RM, Taylor GJ, Parker JDA. The twenty-item Toronto Alexithymia scale—II. Convergent, discriminant, and concurrent validity. *J Psychosom Res.* 1994 Jan; 38(1):33–40. [https://doi.org/10.1016/0022-3999\(94\)90006-x](https://doi.org/10.1016/0022-3999(94)90006-x) PMID: 8126688
87. Moral de la Rubia J, Retamares Rojas R. Estudio de validación de la Escala de Alexitimia de Toronto (TAS-20) en muestra española. [A Toronto Alexithymia Scale (TAS-20) validation study on a Spanish sample]. *Rev Electronica Psicol Psicología.com.* 2000;4(1). Available from: <https://psiquiatria.com/todas/estudio-de-validacion-de-la-escala-de-alexitimia-de-toronto-tas-20-en-muestra-espanola/>
88. Brown KW, Ryan RM. The benefits of being present: mindfulness and its role in psychological well-being. *J Pers Soc Psychol.* 2003 Apr; 84(4):822–48. <https://doi.org/10.1037/0022-3514.84.4.822> PMID: 12703651
89. Carlson LE, Brown KW. Validation of the Mindful Attention Awareness Scale in a cancer population. *J Psychosom Res.* 2005 Jan; 58(1):29–33. <https://doi.org/10.1016/j.jpsychores.2004.04.366> PMID: 15771867
90. Soler J, Tejedor R, Feliu-Soler A, Pascual JC, Martí AC, Palao JS, et al. Propiedades psicométricas de la versión española de la escala Mindful Attention Awareness Scale (MAAS). [Psychometric properties of Spanish version of Mindful Attention Awareness Scale (MAAS)]. *Actas Esp Psiquiatr.* 2012; 40(1):19–26. PMID: 22344492.
91. Lenzenweger MF, Clarkin JF, Kernberg OF, Foelsch PA. The Inventory of Personality Organization: psychometric properties, factorial composition, and criterion relations with affect, aggressive dyscontrol, psychosis proneness, and self-domains in a nonclinical sample. *Psychol Assess.* 2001 Dec; 13(4):577–91. <https://doi.org/10.1037/1040-3590.13.4.577> PMID: 11793901
92. Cosentino S. Validación de la versión española del Inventory of Personality Organization (IPO): contribución al psicodiagnóstico de la organización límite [dissertation]. Barcelona:Universitat Ramon Llull; 2017. Available from: [https://www.tdx.cat/bitstream/handle/10803/404792/Tesi\\_Salvatore\\_Cosentino.pdf?sequence=2&isAllowed=y](https://www.tdx.cat/bitstream/handle/10803/404792/Tesi_Salvatore_Cosentino.pdf?sequence=2&isAllowed=y).
93. Derogatis L. The SCL-90-R clinical psychometric research. Baltimore: Clinical Psychometric Research; 1975.
94. González de Rivera J, Derogatis L, de las Cuevas C, Gracia R, Rodríguez F, Henry M, et al. The Spanish versión of the SCL-90-R. Normative data in general population. Towson: Clinical Psychometric Research; 1989.
95. De las Cuevas C, González de Rivera L, Henry M, Monterrey A, Rodríguez Pulido F, Marco R. Análisis factorial de la versión española del SCL-90-R en la población general. [Factor analysis of the Spanish version of the SCL-90-R in the general population]. *An Psiquiatr.* 1991; 7(3):93–6. Available from: [http://www.psicoter.es/pdf/91\\_A098\\_20.pdf](http://www.psicoter.es/pdf/91_A098_20.pdf).
96. Beck AT, Steer RA, Brown GK. BDI-II. Beck Depression Inventory—Second Edition. Manual. San Antonio (TX): Psychological Corporation; 1996.
97. Sanz J, Navarro M, Vázquez C. Adaptación española del Inventario para la Depresión de Beck-II (BDI-II): 1. Propiedades psicométricas en estudiantes universitarios. [Spanish adaptation of the Beck Depression Inventory-II (BDI-II): 1. Psychometric properties in university students]. *Anal Modif Cond.* 2003; 29:239–88.
98. Sanz J, Perdígón Viñuela A, Vázquez C. Adaptación española del Inventario para la Depresión de Beck-II (BDI-II): 2. Propiedades psicométricas en población general. [Spanish adaptation of the Beck Depression Inventory-II (BDI-II): 2. Psychometric properties in the general population]. *Clin Salud.* 2003 Jan 1; 14:249–80. Available from: <https://journals.copmadrid.org/clysa/art/db8e1af0cb3aca1ae2d0018624204529>.
99. Sanz J, García-Vera MP, Espinosa R, Fortún M, Vázquez C. Adaptación española del Inventario para la Depresión de Beck-II (BDI-II): 3. Propiedades psicométricas en pacientes con trastornos psicológicos. [Spanish adaptation of the Beck Depression Inventory-II (BDI-II): 3. Psychometric properties in patients with psychological disorders]. *Clin Salud.* 2005 Jan 1; 16:121–42. Available from: <https://journals.copmadrid.org/clysa/art/37bc2f75bf1bcfe8450a1a41c200364c>.
100. Krueger RF, Derringer J, Markon KE, Watson D, Skodol AE. Initial construction of a maladaptive personality trait model and inventory for DSM-5. *Psychol Med.* 2012 Sep; 42(9):1879–90. <https://doi.org/10.1017/S0033291711002674> PMID: 22153017
101. Krueger RF, Derringer J, Markon KE, Watson D, Skodol AE. The personality inventory for DSM-5 brief form (PID-5-BF). In: American Psychiatric Association, editor. *Diagnostic and Statistical Manual of Mental Disorders, 5th Edition.* Whashington (DC): American Psychiatric Publishing; 2013.

102. Gutiérrez F, Aluja A, Peri JM, Calvo N, Ferrer M, Baillés E, et al. Psychometric properties of the Spanish PID-5 in a clinical and a community sample. *Assessment*. 2017 Apr; 24(3):326–36. <https://doi.org/10.1177/1073191115606518> PMID: 26391204
103. Goldberg LR. An alternative 'description of personality': the big-five factor structure. *J Pers Soc Psychol*. 1990 Dec; 59(6):1216–29. <https://doi.org/10.1037//0022-3514.59.6.1216> PMID: 2283588
104. Barkham M, Hardy GE, Startup M. The IIP-32: a short version of the Inventory of Interpersonal Problems. *Br J Clin Psychol*. 1996 Feb; 35(1):21–35. <https://doi.org/10.1111/j.2044-8260.1996.tb01159.x> PMID: 8673033
105. Salazar J, Martí V, Soriano S, Beltran M, Adam A. Validity of the Spanish version of the Inventory of Interpersonal Problems and its use for screening personality disorders in clinical practice. *J Personal Disord*. 2010 Aug; 24(4):499–515. <https://doi.org/10.1521/pedi.2010.24.4.499> PMID: 20695809
106. Wyrwich KW, Nienaber NA, Tierney WM, Wolinsky FD. Linking clinical relevance and statistical significance in evaluating intra-individual changes in health-related quality of life. *Med Care*. 1999 May; 37(5):469–78. <https://doi.org/10.1097/00005650-199905000-00006> PMID: 10335749
107. Batista-Foguet JM, Coenders G, Alonso J. Análisis factorial confirmatorio. Su utilidad en la validación de cuestionarios relacionados con la salud. [Confirmatory factor analysis. Its role on the validation of health related questionnaires]. *Med Clin*. 2004; 122 Suppl 1:21–7. <https://doi.org/10.1157/13057542> PMID: 14980156
108. Devins GM, Dion R, Pelletier LG, Shapiro CM, Abbey S, Raiz LR, et al. Structure of lifestyle disruptions in chronic disease: a confirmatory factor analysis of the Illness Intrusiveness Ratings Scale. *Med Care*. 2001 Oct; 39(10):1097–104. <https://doi.org/10.1097/00005650-200110000-00007> PMID: 11567172
109. Hatcher L. A step-by-step approach to using the SAS system for factor analysis and structural equation modeling. Cary (NC): SAS Publishing; 1994.
110. Mulaik SA, James LR, Van Alstine J, Bennett N, Lind S, Stilwell CD. Evaluation of goodness-of-fit indices for structural equation models. *Psychol Bull*. 1989; 105(3):430–45. <https://doi.org/10.1037/0033-2909.105.3.430>
111. Comrey AL, Lee HB. A first course in factor analysis. 2nd ed. New York: Routledge; 2016.
112. Raftery AE. Bayesian model selection in social research. *Sociol methodol*. 1995; 25:111–63. <https://doi.org/10.2307/271063>
113. Cronbach LJ. Coefficient alpha and the internal structure of tests. *Psychometrika*. 1951 Sep 1; 16(3):297–334. <https://doi.org/10.1007/BF02310555>
114. Nunnally JC, Bernstein IH. Psychometric theory. 3rd ed. New York: McGraw-Hill; 1994.
115. Koo TK, Li MY. A Guideline of selecting and reporting intraclass correlation coefficients for reliability research. *J Chiropr Med*. 2016 Jun; 15(2):155–63. <https://doi.org/10.1016/j.jcm.2016.02.012> PMID: 27330520
116. Portney LG, Watkins MP. Foundations of clinical research: applications to practice. Pearson New International Edition. 3rd ed. Upper Saddle River (NJ): Pearson Prentice Hall; 2014.
117. Campbell MJ, Swinscow TDV. Statistics at square one. 11th ed. Chichester (GB): Wiley-Blackwell; 2009. (BMJ Books).
118. Muthén LK, Muthén BO. Mplus User's Guide [Internet]. 8th ed. Angeles Los (CA): Muthén & Muthén; 1998 [cited 2021 Aug 2]. Available from: [https://www.statmodel.com/download/usersguide/MplusUserGuideVer\\_8.pdf](https://www.statmodel.com/download/usersguide/MplusUserGuideVer_8.pdf).
119. Identidad Coderch J., contexto y mentalización. [Identity, context and mentalization]. *Clin Invest Relac*. 2012; 6(2):218–34. Available from: <https://www.psicoterapiarelacional.es/CelRREVISTA-Online/CelR-Buscador-Valore-y-comente-los-trabajos-publicados/ID/303/Identidad-contexto-y-mentalizacion-Joan-Coderch>.
120. Guerini R, Marraffa M, Paloscia C. Mentalization, attachment, and subjective identity. *Front Psychol*. 2015 Jul 21; 6:1022. <https://doi.org/10.3389/fpsyg.2015.01022> PMID: 26257680
121. Kernberg OF. Identity: recent findings and clinical implications. *Psychoanal Q*. 2006 Oct; 75(4):969–1004. <https://doi.org/10.1002/j.2167-4086.2006.tb00065.x> PMID: 17094369
122. Bateman A, Fonagy P. Psychotherapy for borderline personality disorders: mentalization based treatment. Oxford (GB): Oxford University Press; 2004.