Portuguese validation of the credibility/expectancy questionnaire in routine practice

Sara Silva, Eunice Barbosa, João Salgado, Carla Cunha

ISMAI - University Institute of Maia & Center for Psychology at University of Porto, Portugal

ABSTRACT

This study aimed to validate and explore the psychometric properties of the credibility/expectancy questionnaire (CEQ) for the Portuguese population in the context of routine practice. The sample includes 87 clients from a university psychotherapy clinic. All clients completed self-report measures to assess credibility and expectation (CEQ), as well as measures of general mental health, depressive symptoms and therapeutic alliance in session 1 and session 2. The exploratory analysis revealed the existence of two factors in CEQ for the Portuguese population, factor 1 credibility and factor 2 expectation. The measure demonstrated good internal consistency, test-retest reliability and good adjustment in the confirmatory analysis. In terms of convergent validity, no significant correlation was found between credibility and expectation and the therapeutic alliance. The results corroborate the CEQ psychometric qualities for the Portuguese population, also showing its applicability in a context of routine practice.

Key words: Credibility; expectation; validation; routine practice; psychotherapy.

Introduction

Credibility and expectation - a brief construct review

Client's credibility and expectation in psychotherapy have been recently recognized as important aspects for the reduction of clinical symptoms (Constantino, Coyne, Boswell, Iles, & Visla, 2018; Constantino, Visla, Coyne, & Boswell, 2018; Newman & Fisher, 2010). Credibility and expectations refer to the client's beliefs about the treatment and its role in achieving the desired changes, as well as the roles they and their therapists will adopt and engage throughout the therapy (Constantino, Arnkoff, Glass, Ametrano, & Smith, 2011; Constantino, Coyne, *et al.*, 2018; Constantino & Westra, 2012). The recognition of the importance of these aspects is not just based on its impact on symptoms reduction, but also on the impact that clients' perceptions, motivations and actions have on therapy itself, such as a better understanding of the psychotherapeutic process, or a greater adherence and involvement in psychotherapeutic tasks (Constantino *et al.*, 2011; Constantino, Coyne, *et al.*, 2018; Constantino & Westra, 2012). Recently, several authors have been increasingly recommending the integration of empirically validated measures to assess client's credibility and expectation in clinical practice (cf. Andersson, Johansson,

Nordlander, & Asmundson, 2012; Norberg *et al*., 2011; Twohig *et al*., 2010; Wahbeh, Goodrich, & Oken, 2016). However, these measures are often validated for specific intervention context or specific disorders. For this purpose, we will apply the Credibility Expectancy Questionnaire in the context of routine practice in a university psychotherapy clinic and evaluate the psychometric properties of the Portuguese version of this measure.

The credibility/expectancy questionnaire (CEQ) (Devilly & Borkovec, 2000) is one of the mostly used measures to assess the client's credibility and expectancy, which has shown results not only in reducing depressive, anxiety and general symptoms, but also in the increase of the therapeutic alliance in the psychotherapy (cf. Cohen, Beard, & Björgvinsson, 2015; Greenberg, Constantino, & Bruce, 2006; Newman & Fisher, 2010; Sochting, Tsai, & Ogrodniczuk, 2016; Stinson, Perez, Ohrt, Von Schell, & Bruening, 2018; Thompson-Hollands, Bentley, Gallagher, Boswell, & Barlow, 2014; Tompkins, Swift, Rousmaniere, & Whipple, 2017). It is a self-report questionnaire composed by 6 items clustered in two subscales: i) credibility; and ii) outcome expectation (Devilly & Borkovec, 2000). The credibility assesses how reliable the treatment is for the client (Devilly & Borkovec, 2000), which means, how much the client believes he/she can trust that treatment to be helpful for him/her and how helpful it will be. It is often related with the client's perception about the therapist's skills, measuring how much he/she believes in that therapeutic approach (Constantino, Coyne, *et al*., 2018). The expectation refers to the improvement that the client believes will be achieved with the treatment (Devilly & Borkovec, 2000; Stinson *et al*., 2018). The expectation precedes any contact with the therapist and is related to the client's expectation of how much he/she will improve with that particular treatment taking into account his/her perceptions upon it before treatment starts (Constantino, Visla, *et al*., 2018). While the credibility demands a higher cognitive focus, the expectation has a higher focus on emotional processes (Devilly & Borkovec, 2000). For instance, if a specific therapy/treatment was effective in the past with a specific client, the perception would be that it is more credible than other interventions (Constantino, Coyne, *et al*., 2018); or, if a client feels that he/she is not capable to complete a treatment due to a previous experience of dropout, he/she has a lower expectation about the result (Constantino, Visla, *et al*., 2018).

Psychometric properties of credibility/expectancy questionnaire

The psychometric properties of CEQ have been originally explored by Devilly and Borkovec (2000) in three clinical samples across three different contexts. In the first study, the first sample was comprised by 123 participants of a residential program for veterans receiving a one-week intervention focused on improving family functioning, managing depression, anger and anxiety. The CEQ was filled at the day before the beginning of the treatment, after clients were informed about the treatment procedures. In the second study, the second sample was composed by 69 participants with generalized anxiety disorder, allocated into three therapeutic treatment conditions, namely: cognitive therapy (CT), relaxation and self-control desensitization, and also a combination of both. Each participant received 14 treatment sessions and the CEQ measures completed after sessions 1 and 2 were used for test-retest reliability (although the questionnaire was applied at the end of all sessions). Finally, the third study had a sample of 22 participants receiving cognitive behavior therapy (CBT) or eye movement desensitization and reprocessing (EMDR). The CEQ was applied at the end of the assessment session, after the therapeutic rational had been introduced.

Across these samples, the authors found promising reliable results regarding the factorial structure, test-retest, internal consistency, and predictive capacity of CEQ (Devilly & Borkovec, 2000). They found two factors, representing the subscales of credibility and outcome expectation. The internal consistency in each factor was high, ranging from $0.79 < \alpha < 0.90$ for the outcome expectation subscale, and from $0.81 < \alpha < 0.86$ for the credibility subscale. The internal consistency for the complete scale was also high, ranging from $0.84 < \alpha < 0.85$. Also, a high inter-items correlation was found, specifically of $0.53 < r < 0.85$ in the expectation and of $0.62 < r < 0.78$ in the credibility factor. Results suggested a good reliability in the test-retest, namely with significant correlations in the outcome expectation ($r = 0.82$, $P < 0.001$), and in the credibility subscales ($r = 0.75$, $P < 0.001$). Finally, authors also reported that the expectation subscale was a good predictor of outcome ($r = 0.39$, $P < 0.002$) on the generalized anxiety sample (Devilly & Borkovec, 2000).

The CEQ has been validated in other intervention contexts, specifically, to assess parents' beliefs about their child's treatment for conduct problems (Nock, Ferriter, & Holmberg, 2007) and for clients' beliefs on cognitive behavioral intervention for chronic pain (Smeets *et al*., 2008). The results are similar to the original version of the CEQ, showing good psychometric properties and the same factorial structure. Nevertheless, Nock and colleagues (2007) noted that item 5, which aims to assess expectations, cross-loaded on both factors. Even considering that in the original study the item 5 also cross-loaded on the credibility factor, it was suggested that in future studies items should be formulated with 'I think' when assessing credibility and with 'I feel' when tapping expectations.

Coste, Tarquinio, Rouquette, Montel, and Pouchot (2019) recently explored the psychometric properties of CEQ for the French population, namely through confirmatory analysis and test-retest reliability evaluation, with 206 patients diagnosed with post-traumatic stress disorder starting individual therapy. The CEQ was applied before the first session and after the presentation of the therapeutic rationale. The results corroborate those presented in the val-

idation of the original version (Devilly & Borkovec, 2000). Confirmatory factor analysis suggests that the most robust model has two factors, the first of credibility (item1, item 2 and item 3) and the second of outcome expectation (item 4, item 5, item 6; Coste *et al.*, 2019). The test-retest reliability evaluation, presented an excellent internal consistency with α=0.97 for credibility and α=0.95 for expectation (Coste *et al.*, 2019).

Another recent study explored the content validity of CEQ with 17 patients with nonspecific chronic musculoskeletal pain or fibromyalgia from a clinical trial in pain rehabilitation with a motivational interviewing intervention (Mertens, Goossens, Verbunt, Köke, & Smeets, 2013; Mertens, Moser, Verbunt, Smeets, & Goossens, 2017). The main question was how the patients responded and interpreted the CEQ before and during the intervention. The results suggested that there were no significantly differences in understanding the content of CEQ items. However, patients that had finished the treatment denoted the need to change their answer (Mertens *et al.*, 2017). These results indicate that during the treatment, patients present different needs and understand the process of changing expectations in a distinct way (Mertens *et al.*, 2017).

Previous research on the psychometric properties of CEQ has evidenced encouraging results by showing, in addition to good reliability indexes, good adjustment indexes and content validity for this measure. However, convergent and divergent validity was not considered in any of the previous articles (Coste *et al.*, 2019; Devilly & Borkovec, 2000; Mertens *et al.*, 2017; Nock *et al.*, 2007; Smeets *et al.*, 2008; Soto-Pérez & Franco-Martín, 2014), which is a limitation we hope to address in the present study (see below). The original validation article of the CEQ (Devilly & Borkovec, 2000) only explored predictive validity in relation to psychotherapy outcome, taking into account the improvements after the intervention in the depressive symptoms [assessed through the Beck depression inventory-II (BDI-II); Beck, Steer, Brown, 1996] and anxiety symptoms [assessed with the state-trait anxiety inventory (STAI); Spielberger *et al.*, 1970].

Aims

Up until now, according to our review of the published literature there are no measures to assess treatment credibility and outcome expectation adapted to the Portuguese population. Also, previous studies were conducted on small samples (Devilly & Borkovec, 2000; Mertens *et al.*, 2017; Nock *et al.*, 2007; Smeets *et al.*, 2008; Soto-Pérez & Franco-Martín, 2014); therefore, we hope to contribute to the literature on treatment credibility and outcome expectation by expanding the scope of the previous studies and using a larger sample of clients treated in routine psychotherapy practice.

The present study aims to adapt and validate the CEQ (Devilly & Borkovec, 2000) to the Portuguese population, and specifically, to the routine practice of psychotherapy in

a university counseling service. Since the CEQ is one of the credibility and expectancy measures that is most used in research, and previous studies showed good psychometric properties in different contexts of intervention and populations (Coste *et al.*, 2019; Devilly & Borkovec, 2000; Mertens *et al.*, 2017; Nock *et al.*, 2007; Smeets *et al.*, 2008; Soto-Pérez & Franco-Martín, 2014), the development of a Portuguese validation will contribute to foster research in this field. We aim to explore the psychometric properties and the factorial structure of CEQ, namely through exploratory factor analysis, confirmatory factor analysis, reliability test, test-retest. We will also expand the scope of the previous validation studies by analyzing the discriminant and convergent validity of the CEQ with depressive symptoms (assessed through the BDI; Beck, Steer, & Brown, 1996), anxiety symptoms (through the STAI; Spielberger *et al.*, 1970), general symptoms [clinical outcome routine evaluation - outcome measure 10 (CORE-10); Barkham *et al.*, 2013] and the therapeutic alliance [working alliance inventory-6 (WAI-6) items; Falkenström, Hatcher, Skjulsvik, Larsson, & Holmqvist, 2015].

Methods

Participants

This study involved 87 adult participants, all psychotherapy clients in a university counseling service, of whom 56.3% (n=49) were women and 43.7 % (n=38) were men, with no statistical differences found between these groups ($\chi^2(86)$=87.00, P=0.450). Participants' ages ranged from 18 to 62 years old, with a mean age of 28 years [standard deviation (SD)=10.25]; 44.8% (n=39) of the participants were employed, 36.8% were students (n=32); 9.2% (n=8) were unemployed; 6.9% (n=6) were working-students and 2.2% (n=2) were maids or held other jobs from home. Also, 85.1% (n=74) of the participants were single, 10.3 (n=9) were married or in co-habitation and 4.6% (n=4) were divorced.

Routine practice

Participants were recruited from a university counseling service (University Institute of Maia (ISMAI), Portugal) with the aim to provide psychological support, psychotherapy and/or personal development services to the university community and to the surrounding community in general. Even though not all diagnostic information was available to all clients of this sample, to the ones that were, most were exhibiting the more common and prevalent emotional disorders, such as depression, anxiety and related disorders (Timulak *et al.*, 2020). Therapists integrated in this counseling service adopt different empirically validated therapeutic approaches (*e.g.*, emotion focused therapy, cognitive behavioral therapy) and receive clinical supervision provided by senior therapists, to support their evidence based clinical practices. As mentioned, the psychotherapy inter-

ventions follow empirical validated protocols ranging from 16 to 20 sessions. All the participants undergo an evaluation session, which implies the application of a clinical interview according with Diagnostic and Statistical Manual of Mental Disorders - 5th (DSM-V) (American Psychiatric Association, 2013), namely *The Structured Clinical Interview for DSM-5® Disorders: Clinician Version* (SCID-V-CV) (First, Williams, Karg, & Spitzer, 2016), and the application of several questionnaires to assess depressive, anxious and general clinical symptoms. Along the psychotherapy process, these clinical symptoms are also monitored to provide an ongoing monitoring of client change throughout the therapeutic process.

Measures

The SCID-V-CV (First *et al.*, 2016) is a semi-structured interview which allows to collect, at an initial assessment moment, general data about psychopathological symptoms and disorders throughout the life cycle taking into account the classification and diagnostic criteria of the DSM-V (American Psychiatric Association, 2013).

The CEQ (Devilly & Borkovec, 2000) is a self-report measure with 6 items to assess the clients' treatment credibility and outcome expectancy. In the credibility factor the items explore: if the treatment appears logical; if the treatment appears useful; and if the treatment is reliable. The outcome expectation items explore: the improvement that the patient thinks will probably occur; if the patient feels that this therapy will help him/her; if the patient feels that improvement will occur (Coste *et al.*, 2019). This measure is scored on a Likert scale ranging from 1 to 9 (*e.g.*, 1 means 'not at all', 5 'somewhat', and 9 'very much') or from 0% to 100% (in 10-point increments). Items 4 and 6, coded from 0%-100%, are converted linearly on a Likert scale from 1 to 9 (Devilly & Borkovec, 2000; Nock *et al.*, 2007; Smeets *et al.*, 2008). The CEQ in the original version is composed by two subscales: i) treatment credibility (items 1, 2, and 3); and ii) outcome expectation (items 4, 5, and 6). The final scoring of this measure delivers a subscale score (computed by the sum of the items that compose each subscale), with a higher quotation indicating greater treatment credibility and outcome expectation (Devilly & Borkovec, 2000; Nock *et al.*, 2007; Smeets *et al.*, 2008). As previously mentioned, the original English version of CEQ showed good psychometric properties with an internal consistency ranging from $0.84 < \alpha < 0.85$.

The BDI-II (Beck, Steer, & Brown, 1996) is a self-report questionnaire that assesses the intensity of depressive symptoms. The Portuguese version of the BDI-II has been developed by Coelho, Martins, and Barros (2002) from Beck and colleagues (1996). This measure comprises 21 items scored in a Likert scale ranging from 0 to 3 points. The total score indicates the intensity of depressive symptoms: mild depressive symptoms (from 0 to 13 points), mild depression (from 14 to 19 points), moderate depression (20

to 28 points) and severe depression (from 29 to 63 points). The Portuguese version has an internal consistency of 0.91 (Steer, Brown, Beck, & Sanderson, 2001).

The STAI (Spielberger *et al.*, 1970) is a self-report questionnaire that evaluates the intensity of anxiety symptoms. The STAI is comprised by two sets of 20 statements to assess both the temporary, transitory anxiety (STAI-State) and the general anxiety (STAI-Trait). This instrument is scored in a Likert scale of 1 to 4 points. The Portuguese version presented good internal consistency, both on the STAI-State (a=0.89) and STAI-Trait (a>0.88; Silva & Campos, 1998; Silva, 2006).

The CORE-10 (Barkham *et al.*, 2013) is a self-report measure that assesses mental health symptoms in different contexts. The CORE-10 is composed by 10 items clustered into 4 subscales: subjective well-being; complaints and symptoms; social and personal functioning; and risk behavior (Barkham *et al.*, 2013). The 10 items are scored on a Likert scale that ranges from 0 ('never') to 4 ('always'). The CORE-10 is a reduced version of the *clinical outcome routine evaluation - outcome measure* (CORE-OM) (Sales, Moleiro, Evans, & Alves, 2012), a 34 items questionnaire translated and adapted for the Portuguese population (Sales *et al.*, 2012). The Portuguese version presented a good internal consistency ($\alpha > 0.80$).

The WAI-6 (Falkenström *et al.*, 2015) is composed of 6 items that evaluate the quality of the therapeutic relationship using a likert scale from 1 ('never') to 5 ('always'; Falkenström *et al.*, 2015). The WAI-6 is a reduced version of the 12 items *working alliance inventory - short revised* (WAI-SR) (Horvath & Greenberg, 1989) translated and validated to the Portuguese population by Machado and Horvath (1999). The Portuguese version has an adequate reliability ($\alpha=0.95$; Machado & Horvath, 1999).

Procedures

First, authorization was requested and provided by the authors of the CEQ for the translation, use, and validation of the measure for the Portuguese population. Second, the measure was translated into Portuguese. The translation process went through different phases: initially, the translation was carried out by two independent researchers who speak both languages; then, a synthesis of the first translation was carried out by the main researcher; finally, there was a back-translation and revision performed by an auditor, fluent in both languages. Third, the current study was submitted and approved by the ethical committee of the University (ISMAI, Portugal). The goals of the research were initially presented to all clients initiating psychotherapy at the university psychotherapy clinic and, when they agreed to participate in this study, researchers gathered their informed consent regarding the collection of the assessment data. Following the protocol of the clinic, all the clients had an initial assessment with the SCID-V, in order to assess clinical diagnosis. These clinical data are included in the clinical and confidential processes managed by the thera-

pists, and the informed consent approved for this study did not include access to them. In this study only anonymous data collected a priori. Also, in this session where filled the BDI-II and the STAI measures. The CEQ was applied in at the end of session 1 and 2 (with a two week interval between the sessions). The CORE-10 was applied at the beginning of sessions 1, 2 and 3; and the WAI-6 at the end of the session 1 and 2. This assessment protocol allowed the analysis of the discriminant (BDI, STAI, CORE-10), convergent (WAI) and predictive (CORE-10) validity of CEQ. The repeated application of the CEQ allowed for the test-retest reliability analysis.

Statistical analysis

For statistical analysis we used the SPSS 25 and AMOS 24 (SPSS Inc., Chicago, IL, USA). An exploratory factor analysis (EFA) was performed following the procedures of the original study (Devilly & Borkovec, 2000). Factor extraction was carried out by the *principal axis factoring* method followed by an oblique rotation, since the underlying factors (*i.e.*, treatment credibility and outcome expectation) are assumed to be related to each other (Field, 2017). According to the Kaiser criterion, factors with an eigenvalue of greater than 1 were retained (Fabrigar & Wegener, 2012; Field,2017; Kaiser,1960; Marôco, 2018). The internal consistency of each factor was evaluated by the Cronbach alpha coefficient (0.60 to 0.70 acceptable and >0.70 recommended; Field, 2017; Marôco, 2018).

Also, factorial confirmatory analysis (CFA) was performed following the similar strategy adopted by Kaya *et al.* (2015) and Van Prooijen and Van der Kloot (2001). The adjustment quality of the factorial model was evaluated according to the indexes with empirical statistical support (Brown, 2015; Flora, 2018; Kline, 2016; Marôco, 2014), specifically: Chi-square of adjustment (χ^2/df); goodness of fit index (GFI >0.90); comparative fit index (CFI >0.90); root mean square error of approximation (RMSEA).

We also performed a convergent and discriminant validity analysis. The convergent validity refers to the analysis of scales constituting the same or an identical construct and, therefore, it is expected that these measures present positive and high correlations between them (Marôco, 2014). In this study, we explored the correlation between CEQ and WAI-C. The discriminant validity refers to the correlation with other measures that evaluate different constructs, and therefore, expecting low correlations between these (Marôco, 2014). For this kind of analysis, we carried out a Pearson correlation with measures of depressive symptoms (BDI), anxious symptoms (STAI) of session 0, general symptoms of session 1 (CORE-10) with the therapeutic alliance CEQ. The predictive validity was performed with CEQ in session 1 and with CORE-10 in session 3 following the procedures of regression models.

Results

Factor structure

The descriptive statistics and tests for the normality of the sample were conducted among the items to ensure no violations of the techniques assumptions were evident. The Kaiser-Meyer-Olkin measure verified the sample adequacy for the analysis (KMO=0.73). According to several authors, the Kaiser-Meyer-Olkin >70 is acceptable (Field, 2017; Marôco, 2018). The item inter-correlation matrix showed a number of significant correlation and suggest that the questionnaire has a structure which could be detected by factor analysis [χ^2 (15)=246.636 P<0.000]. The power to conduct the principal axis factoring was achieved. The initial statistic indicated the presence of 2 factors with eigenvalues above 1, which explains 75,5% of the total variance (see Table 1).

According to the theory, treatment credibility and outcome expectation are related variables. In fact, for this sample, the two factors derived a correlation of $r= -0.39$; thus, an oblique rotation was performed. As it can been seen in Table 2, the communalities indicate that the extracted factors accounted for between 0.49% and 0.94%

Table 1. Factor descriptive analysis of Portuguese version of credibility/expectancy questionnaire.

Factor	Eigen value	% of variance	Cumulative %
1	3269	54,479	54,479
2	1261	21,009	75,488

Table 2. Credibility/expectancy questionnaire factor structure for Portuguese population.

Item	Factor 1	Factor 2	Comm	Mean	σ
1. How logical does the therapy offered to you seem	0.72	−0.1	0.56	8.38	0.7
2. How successful do you think this treatment will be in reducing your symptoms	0.67	−0.45	0.49	8.09	0.93
3. How confident would you be in recommending this treatment to a friend	0.77	−0.42	0.61	8.57	0.62
4. How much improvement in your symptoms do you think will occur	0.37	−0.84	0.58	7.79	0.98
5. How much do you really feel that therapy will help you to reduce your symptoms	0.68	0.59	0.71	8.24	0.79
6. How much improvement in your symptoms do you really feel will occur	0.35	−0.97	0.94	7.62	1.2

of that item's variance, with items 1, 2, 3 loading on factor 1 (credibility) and items 4 and 6 loading on factor 2 (expectation). The item 5 is cross-loaded in both factors. Due to this, we followed the recommendation of Costello and Osborne (2005) for the removal of this item, since it does not give a clear explanation of the constructs and potentially impoverishes them.

Reliability

We analyzed each factor in terms of internal consistency to derive the Cronbach's a. We also analyzed the item's total correlation, without item 5 (removed according to Costello & Osborne, 2005). A Cronbach's a=0.75 was found for the treatment credibility factor, and a Cronbach's a=0.90 was found for the outcome expectation factor. A Cronbach's a=0.78 was found for the complete scale.

We assessed test-retest reliability by correlating the factor scores at session 1 with the scores at session 2. In the credibility factor, the results of the test-retest reliability were significant, r (52)=0.44, P=0.001; in the expectation factor, the results also showed a significant correlation, r (51)=0.61, P=0.000. When both factors were combined, the test-retest reliability was significant, r (51)=0.59, P=0.00.

Confirmatory analyses

Confirmatory factor analysis was performed with the same sample in order to test the stability of the exploratory factor analysis. Initially, the adjustment of the model was made from the modification indexes (higher than 4; P<0.001) produced by AMOS. The model followed the factorial structure of the exploratory phase with the distribution of two factors, one of treatment credibility (item 1, item 2, and item 3) and other of outcome expectation (item 4 and item 6) that showed good adequacy of the items, since none of the items had a factor load of less than 0.50 (see Figure 1). However, a good model fit was not found {χ^2 (4)=13.711; χ^2/df=3.428 P=0.008, CFI=0.94, GFI=0.95, TIL=0.86, RMSEA=0.17}. Thus, we analyzed the change indexes in the model, suggesting a cross-load of the error from item 2 to item 3 were analyzed. According to Brown (2015), a small modification in the indexes can produce a good-fitting model improving the χ^2 with 1 df. In this sense, the modification index was changed from 4 to 1.

After model re-adjustment, the results of the confirmatory factor analysis indicate an appropriate adjustment model {χ^2 (3)=0.596; χ^2/df=0.199 P=0.897, CFI=1.00, GFI=0.99, TIL=1.05, RMSEA=0.000} with the existence of two factors: the credibility factor constituted by item 1 (λ=0.17), item 2 (λ=0.65) and item 3 (λ=1.00); and the expectation factor constituted by item 4 (λ=1.00) and item 6 (λ=1.13), respectively (see Figure 2). It should be noted that there is a cross-load of the error of item 3 with the error of item 2.

Convergent validity

Convergent validity analysis was performed between the CEQ and WAI-6 items. Results evidenced that the association between treatment credibility on session 1 and therapeutic alliance on session 2 was not statistically significant, r (68)=0.23, P=0.056. Similarly, the association between outcome expectation and the therapeutic alliance was not statistically significant, r (68)=0.21, P=0.09.
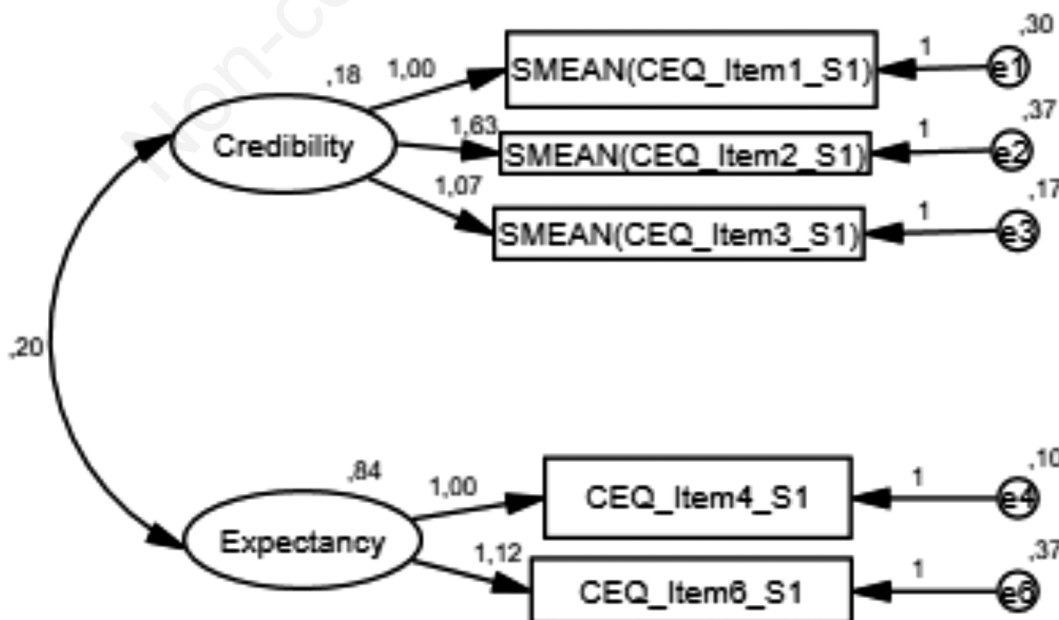

Figure 1. Path diagrams of confirmatory model of credibility/expectancy questionnaire (CEQ) with for two factor model without a good fit adjustment's indices {χ^2 (4)=13.711; χ^2/df=3.428 P=0.008, CFI=0.94, GFI=0.95, TIL=0.86, RMSEA=0.17}.

Discriminant validity

Regarding the analysis of the discriminant validity, we considered the CEQ of the first session and the measurement of symptoms in the assessment session and session 1, namely through the BDI-II, STAI-T, STAI-S and CORE-10. No statistically significant associations were found (see Table 3).

Predictive validity

In terms of predictive validity, results show an association between the treatment credibility and the outcome ex-

pectation levels reported in session 1, with the symptoms assessed in session 3, as measured by CORE-10. More specifically, results indicate that the outcome expectation of session 1 is significantly associated with a decrease of symptoms of session 3, $r= -0.30$, P=0.019. Regarding credibility, there was no significant association with symptoms of the session 3, $r= -0.12$, P=0.372. The predictive validity of CEQ and symptoms at session 3 was also examined through a linear regression analysis (see Table 4). The results indicate that the outcome expectation significantly predicts the decrease of symptoms in session 3 ($b= -30$; $t(60)= -2.41$; P=0.019). However, outcome expectation

Table 3. Correlation of credibility/expectancy questionnaire and symptoms measures: discriminant validity.

	BDI		STAI-T		STAI-S		CORE-10	
	r_p	P	r_p	P	r_p	P	r_p	P
Credibility	0.073	0.526	−0.043	0.691	0.097	0.370	−0.179	0.097
Expectancy	−0.012	0.916	−0.154	0.154	−0.124	0.252	−0.176	0.103

BDI, Beck depression inventory; STAI-T, state-trait anxiety inventory-trait; STAI-S, state-trait anxiety inventory-state; CORE-10, clinical outcome routine evaluation - outcome measure 10.

Table 4. Linear regression analysis of credibility/expectancy questionnaire and general symptoms (CORE-10) in session 3.

General symptoms	R^2_a	B (SE; CI 95%)	β	t (df)	P
Expectation	0.074	−0.90 (0.37; [−1.65-1.52])	−0.30	−2.41(60)	0.019

SE, standard error; CI, confidence interval.
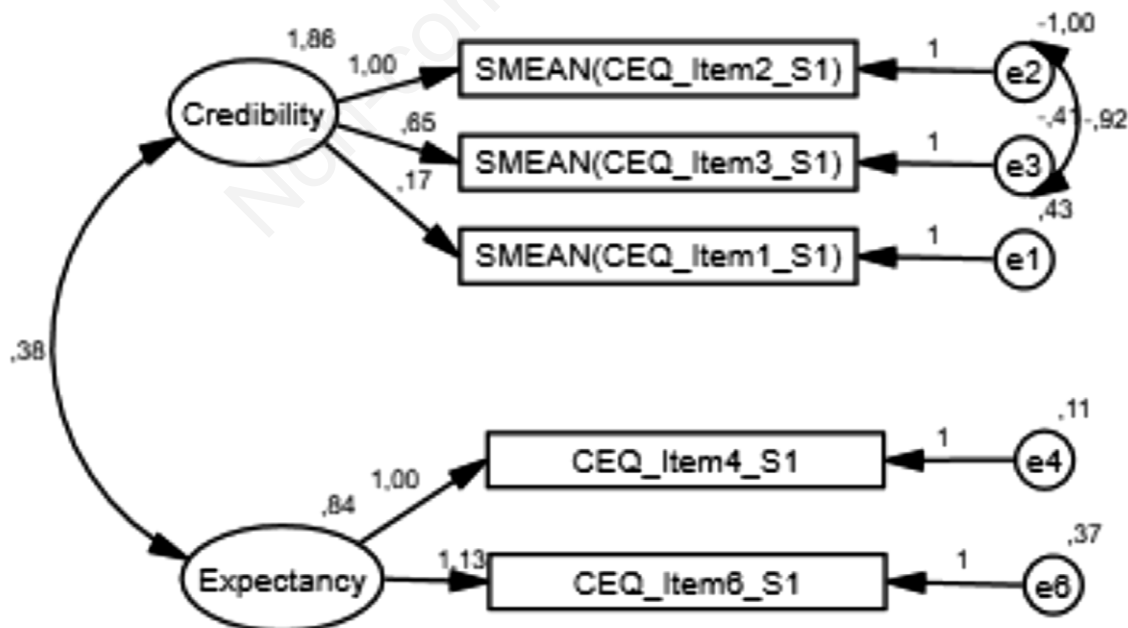
Figure 2. Path diagrams of confirmatory model of credibility/expectancy questionnaire (CEQ) with for two factor model, factor of credibility, with item 1, item 2 and item 3, with a cross-load of the error of item 3 on the error of the item 2; and the factor of expectation with item 4 and item 6. This model presented an adequate fit adjustment's indices {χ^2 (3)=0.596; χ^2/df=0.199 P=0.897, CFI=1.00, GFI=0.99, TIL=1.05, RMSEA=0.000}.

seems to have a low effect on the variability of decreased clinical symptoms.

Discussion

This research aimed to explore the psychometric properties of the CEQ for the Portuguese population in routine clinical practice. We investigated the factorial structure (exploratory and confirmatory), internal consistency, test-retest, convergent, discriminant validity and the predictive validity of the scale in decreasing symptoms.

The results suggest the presence of two factors, one of treatment credibility and one of outcome expectancy, similar to what was found in the original validation study (Devilly & Borkovec, 2000), as well as in other similar studies (*e.g.*, Coste *et al*., 2019; Mertens *et al*., 2017; Nock *et al*., 2007; Smeets *et al*., 2008). However, in previous exploratory factor analysis, item 5 presented a cross-load factor in both factors (Nock *et al*., 2007), which is a result that was also found in the present study. In these cases, several authors recommend the removal of the item, since it does not give a clear explanation of the constructs and impoverishes them (Costello & Osborne, 2005). This is what we opted to do in the present Portuguese version of the CEQ. The cross-load of item 5 in both factors may be due to the formulation of the item, and it is consistent with previous findings in other studies. Actually, the empirical results support the view that this item ('At this point, how much do you feel this treatment will help you reduce your symptoms?') combines elements related both with treatment credibility (how helpful do you think this treatment will be?) and outcome expectations (what do you feel and expect out of this treatment?). In sum, the cross-load item 5 in both factors for the Portuguese version suggests the need of future studies searching for items assessing expectations while avoiding this overlap with the other subscale.

Regarding internal consistency, results show that, after excluding item 5, the Portuguese version of CEQ presents a high internal consistency for each factor (treatment credibility a=0.75 and outcome expectation a=0.90), as well as for the full scale (a=0.78; Field, 2017; Marôco, 2018). Similarly, a test-retest analysis with session 2 also shows good reliability, even without the item 5 (Field, 2017; Marôco, 2018). These results are consistent with the results obtained in other investigations, being a good indicator of the psychometric properties of the CEQ (Coste *et al*., 2019; Devilly & Borkovec, 2000; Nock *et al*., 2007).

In the convergent validity, the association between the CEQ credibility factor and the therapeutic alliance in session 2 was not statistically significant in the present study (Marôco, 2014). However, previous research suggests that credibility and expectation are associated with the therapeutic relationship and these processes may influenced each other throughout therapy (Ametrano, 2011; Smeets *et al*., 2008; Sochting *et al*., 2016). It is possible, though, that the

present results may be explained by the fact that the alliance was considered here (assessed through the WAI) during the initial sessions, when the relationship is still being established (Ardito & Rabellino, 2011). Furthermore, research on credibility and expectations is still recent, and future studies should probably improve the criteria for assessing convergent validity, namely with a better construct than the therapeutic alliance.

Regarding discriminant validity, no significant association was found between CEQ and clinical symptoms. Prior research on this issue is not consistent; however, there are several previous studies that also did not find a correlation between the symptoms and the credibility and expectation (Devilly & Borkovec, 2000; Smeets *et al*., 2008; Sochting *et al*., 2016; Thompson-Hollands *et al*., 2014). Nevertheless, the absence of a significant correlation between CEQ and clinical measures targeting mental health symptoms are within the expected results of a discriminant analysis, giving further support to the use of a 5-item version in the Portuguese context.

Finally, regarding the predictive validity of CEQ and symptomatic decrease in session 3, our results suggest that outcome expectation, specifically, has an effect in decreasing symptoms. These results are consistent with those of other validations (Cohen *et al*., 2015; Devilly & Borkovec, 2000; Smeets *et al*., 2008), which evidence that outcome expectation predict the decrease of the depressive and anxiety symptoms, treatment satisfaction and the improvement of well-being.

Previous studies on the confirmatory analysis of CEQ (Coste *et al*., 2019; Smeets *et al*., 2008) confirm the two-factor model, one for treatment credibility and one for outcome expectation. The present study also presented a similar and adjusted confirmatory model.

The results of this study, in addition to consolidating the psychometric properties of CEQ, also suggest its applicability in the context of routine practice. This validation for routine psychotherapy practice was a first step, which may later enable researchers to better understand how clients' characteristics may influence their beliefs in the therapeutic process. Recently, specific research suggests that some clinical characteristics of the participants may be related to treatment credibility and outcome expectations, such as the diagnosis, medication and previous experiences with psychotherapy (Constantino, Visla, *et al.*, 2018), which were not analyzed here, but should be considered in future investigations. Despite these limitations, the present study contributes to this field by expanding the applicability of this measure, allowing for replicability in a larger sample than the previous ones used for the exploratory factor analysis (*e.g.*, Coste *et al*., 2019; Mertens *et al*., 2017; Nock *et al*., 2007; Smeets *et al*., 2008). However, there are some limitations in the present study that may influence the interpretation of the results presented here, specifically related to divergent and convergent validity (*i.e.*, non-significant results), and with the

confirmatory factor analysis (performed with a small sample, which was the same sample used in the exploratory factor analyzes).

In the future, and following the recommendation of Devilly and Borkovec (2000), we aim to explore the differences between the outcome expectation, contrasting a higher focus on emotional processes *versus* a cognitive focus. Possibly, this will also help to clarify the role of treatment credibility and outcome expectations in predicting therapeutic outcome (good *versus* poor outcome) and psychotherapy dropout. Therefore, an emphasis on longitudinal and process-outcome research is required, namely, to understand how treatment credibility and outcome expectations may evolve throughout psychotherapy and clarify their specific contributions to clients' involvement during the therapeutic process.

References

Ardito, R., & Rabellino, D. (2011). Therapeutic alliance and outcome of psychotherapy: historical excursus, measurements, and prospects for research. *Frontiers in Psychology, 2*, 1-11. doi: 10.3389/fpsyg.2011.00270

American Psychiatric Association. (2013). *Diagnostic and statistical manual of mental disorders* (5th ed.). Arlington, VA: American Psychiatric Association.

Ametrano, R. (2011). Patient outcome expectations and credibility beliefs as predictors of the alliance and treatment outcome. *Masters Theses*, *February*. Retrieved from: http://scholarworks.umass.edu/theses/658

Andersson, G., Johansson, C., Nordlander, A., & Asmundson, G. J. G. (2012). Chronic pain in older adults: A controlled pilot trial of a brief cognitive-behavioural group treatment. *Behavioural and Cognitive Psychotherapy*, *40*(2), 239-244. https://doi.org/10.1017/S1352465811000646

Barkham, M., Bewick, B., Mullin, T., Gilbody, S., Connell, J., Cahill, J., Mellor-Clark, J., Richards, D., Unsworth, G., & Evans, C. (2013). The CORE-10: A short measure of psychological distress for routine use in the psychological therapies. *Counselling and Psychotherapy Research*, *13*(1), 3-13. https://doi.org/10.1080/14733145.2012.729069

Brown, T. (2015). Confirmatory factor analysis for applied research. In *The American Statistician* (Vol. 62, Issue 1). https://doi.org/10.1198/tas.2008.s98

Cohen, M., Beard, C., & Björgvinsson, T. (2015). Examining patient characteristics as predictors of patient beliefs about treatment credibility and expectancies for treatment outcome. *Journal of Psychotherapy Integration*, *25*(2), 90-99. https://doi.org/10.1037/a0038878

Constantino, M. J., Arnkoff, D. B., Glass, C. R., Ametrano, R. M., & Smith, J. Z. (2011). Expectations. *Journal of Clinical Psychology*, *67*(2), 184-192. https://doi.org/10.1002/jclp.20754

Constantino, M. J., Coyne, A. E., Boswell, J. F., Iles, B. R., & Visla, A. (2018). A meta-analysis of the association between patients' early perception of treatment credibility and their posttreatment outcomes. *Psychotherapy*, *55*(4), 486-495. https://doi.org/10.1037/pst0000168

Constantino, M. J., Visla, A., Coyne, A. E., & Boswell, J. F. (2018). A meta-analysis of the association between patients' early treatment outcome expectation and their posttreatment

outcomes. *Psychotherapy*, *55*(4), 473-485. https://doi.org/10.1037/pst0000169

Constantino, M. J., & Westra, H. A. (2012). An expectancy-based approach to facilitating corrective experiences in psychotherapy. *Transformation in Psychotherapy: Corrective Experiences across Cognitive Behavioral, Humanistic, and Psychodynamic Approaches*, 121-139. https://doi.org/10.1037/13747-008

Coste, J., Tarquinio, C., Rouquette, A., Montel, S., & Pouchot, J. (2019). Cross-cultural adaptation and validation of the French version of the credibility/expectancy questionnaire. Further insights into the measured concepts and their relationships. *Psychologie Francaise*, 1-12. https://doi.org/10.1016/j.psfr.2018.11.001

Costello, A. B., & Osborne, J. W. (2005). Best practices in exploratory factor analysis: Four recommendations for getting the most from your analysis. *Practical Assessment, Research and Evaluation*, *10*(7).

Devilly, G. J., & Borkovec, T. D. (2000). Psychometric properties of the credibility/expectancy questionnaire. *Journal of Behavior Therapy and Experimental Psychiatry*, *31*(2), 73-86. https://doi.org/10.1016/S0005-7916(00)00012-4

Fabrigar, L. & Wegener, D. (2012). Exploratory factor analysis. New York, NY: Oxford University Press.

Falkenström, F., Hatcher, R. L., Skjulsvik, T., Larsson, M. H., & Holmqvist, R. (2015). Development and validation of a 6-item working alliance questionnaire for repeated administrations during psychotherapy. *Psychological Assessment*, *27*(1), 169-183. https://doi.org/10.1037/pas0000038

Field, A. (2017). Discovering Statistics Using IBM SPSS Statistics (5th ed). London: SAGE Publications.

First, M.B., Williams, J.B.W., Karg, R.S., & Spitzer, R.L. (2016). Structured clinical interview for DSM-5 disorders, clinician version (SCID-5-CV). Arlington, VA: American Psychiatric Association.

Greenberg, R. P., Constantino, M. J., & Bruce, N. (2006). Are patient expectations still relevant for psychotherapy process and outcome? *Clinical Psychology Review*, *26*(6), 657-678. https://doi.org/10.1016/j.cpr.2005.03.002

Kaya, C., Tansey, T., Chan, F., Bezyak, J., Melekoğlu, M., Çakıroğlu, O., & Köse, S. (2015). Dimensionality of the Turkish version of the self-stigma of seeking help scale: results from exploratory and confirmatory factor analyses. *International Journal for the Advancement of Counselling*, 37:105-116. doi: 10.1007/s10447-014-9230-3

Kline, R. B. (2016). Principles and practices of structural equation modelling (4th Ed.). New York, NY: The Guilford Press.

Mertens, V. C., Goossens, M. E. J. B., Verbunt, J. A., Köke, A. J., & Smeets, R. J. E. M. (2013). Effects of nurse-led motivational interviewing of patients with chronic musculoskeletal pain in preparation of rehabilitation treatment (PREPARE) on societal participation, attendance level, and cost-effectiveness: Study protocol for a randomized controll. *Trials*, *14*(1). https://doi.org/10.1186/1745-6215-14-90

Mertens, V. C., Moser, A., Verbunt, J., Smeets, R., & Goossens, M. (2017). Content validity of the credibility and expectancy questionnaire in a pain rehabilitation setting. *Pain Practice*, *17*(7), 902-913. https://doi.org/10.1016/j.cognition.2017.03.009

Newman, M. G., & Fisher, A. J. (2010). Expectancy/credibility change as a mediator of cognitive behavioral therapy for generalized anxiety disorder: mechanism of action or proxy for symptom change? *International Journal of Cognitive Ther-*

apy, *3*, 245-261. https://doi.org/10.1521/ijct.2010.3.3.245.Expectancy/Credibility

Marôco, J. (2014). *Análise de equações estruturais: Fundamentos teóricos, software e aplicações* (2nd Ed). Pêro Pinheiro, Portugal: ReportNumber.

Marôco, J. (2018). *Análise estatística com o SPSS Statistics* (6ª ed). Pêro Pinheiro, Portugal: ReportNumber.

Nock, M. K., Ferriter, C., & Holmberg, E. (2007). Parent beliefs about treatment credibility and effectiveness: assessment and relation to subsequent treatment participation. *Journal of Child and Family Studies*, *16*(1), 27-38. https://doi.org/10.1007/s10826-006-9064-7

Sales, C. M. D., Moleiro, C. M. de M., Evans, C., & Alves, P. C. G. (2012). Versão Portuguesa do CORE-OM: tradução, adaptação e estudo preliminar das suas propriedades psicométricas TT - The Portuguese version of the CORE-OM: translation, adaptation, and preliminary study of psychometric properties. *Revista de Psiquiatria Clínica (São Paulo)*, *39*(2), 54-59. https://doi.org/10.1590/S0101-60832012000200003

Smeets, R. J. E. M., Beelen, S., Goossens, M. E. J. B., Schouten, E. G. W., Knottnerus, J. A., & Vlaeyen, J. W. S. (2008). Treatment expectancy and credibility are associated with the outcome of both physical and cognitive-behavioral treatment in chronic low back pain. *Clinical Journal of Pain*, *24*(4), 305-315. https://doi.org/10.1097/AJP.0b013e318164aa75

Sochting, I., Tsai, M., & Ogrodniczuk, J. (2016). Patients' perceptions of treatment credibility and their relation to the outcome of group CBT for depression. *Archives of Psychiatry and Psychotherapy*, *18*(4), 7-15. https://doi.org/10.12740/APP/66485

Soto-Pérez, F., & Franco-Martín, M. (2014). PsicoED: Una alternativa online y comunitaria para la psicoeducación en esquizofrenia. *Psicoperspectivas*, *13*(3), 118-129. https://doi.org/10.5027/PSICOPERSPECTIVAS-VOL13-ISSUE3-FULLTEXT-416

Stinson, E. J., Perez, M., Ohrt, T. K., Von Schell, A., & Bruening,

A. B. (2018). The association between program credibility, expectancy, and acceptability with baseline pathology and outcome for a body acceptance prevention program. *Journal of Clinical Psychology*, *74*(12), 2161-2172. https://doi.org/10.1002/jclp.22654

Thompson-Hollands, J., Bentley, K., Gallagher, M., Boswell, J., & Barlow, D. (2014). Credibility and outcome expectancy in the unified protocol: relationship to outcomes. *Journal of Experimental Psychopathology*, *5*(1), 72-82. https://doi.org/10.5127/jep.033712

Timulak, L, Keogh, D, McElvaney, J., Schmitt, S., Hession, N., Timulakova, K., Jennings, C., & Ward, F. (2020). Emotion-focused therapy as a transdiagnostic treatment for depression, anxiety and related disorders: Protocol for an initial feasibility randomised control trial. *HRB Open Research*, *3*, 7. doi.org/10.12688/hrbopenres.12993.1

Tompkins, K. A., Swift, J. K., Rousmaniere, T. G., & Whipple, J. L. (2017). The relationship between clients' depression etiological beliefs and psychotherapy orientation preferences, expectations, and credibility beliefs. *Psychotherapy*, *54*(2), 201-206. https://doi.org/10.1037/pst0000070

Twohig, M. P., Hayes, S. C., Plumb, J. C., Pruitt, L. D., Collins, A. B., Hazlett-Stevens, H., & Woidneck, M. (2010). A randomized clinical trial of acceptance and commitment therapy vs. progressive relaxation training for obsessive compulsive disorder. *Journal of Consulting and Clinical Psychology*, *78*(5), 705-716.

Van Prooijen, J. W., & Van der Kloot, W. A. (2001). Confirmatory analysis of exploratively obtained factor structures. *Educational and Psychological Measurement*, *6*, 777-792.

Wahbeh, H., Goodrich, E., & Oken, B. S. (2016). Internet-based mindfulness meditation for cognition and mood in older adults: a pilot study. *Alternative Therapies in Health and Medicine*, *22*(2), 44–53.