Psychometric properties of the Brazilian version of Clinical Perfectionism Questionnaire after rewording negatively keyed items

Luiz Fellipe Dias da Rocha (iD) * / José Augusto Evangelho Hernandez (iD) * / Eliane Mary de Oliveira Falcone (iD) *

* Universidade do Estado do Rio de Janeiro, Rio de Janeiro, Brasil

Abstract: The present study aims to verify the psychometric properties of the Brazilian version of Clinical Perfectionism Questionnaire after rewording negatively keyed items into positive (CPQ+). It was used a convenience sample of 168 Brazilian university students (*M*=25.37, *SD*=7.40), being 76.8% women. After rewording the negatively keyed items into positive, the CPQ+ was applied. The Depression, Anxiety and Stress Scales – Short Form (DASS-21) were also applied. The Corrected Item-Total Correlation values of CPQ+ items were close to or higher than the cutoff point. After exclusion of item 8, due to crossloading, and correlation between errors of items 3 and 10, the correlated two-factor model presented the best fit of latent structure to CPQ+, through Confirmatory Factor Analysis. The Perfectionistic Concerns dimension of CPQ+ was greater related to global score of DASS-21 in comparison to the Perfectionistic Striving dimension, indicating good validity regarding other variables. In summary, results recommend the use of CPQ+ presented in this study, it is, a version with 11 positive items.

Keywords: Personality, Perfectionism, Measurement, Psychometrics, Psychopathology.

Perfectionism might be defined as a personality trait characterized by setting high personal demanding standards and striving for flawlessness (Stoeber, 2016). Although initial conceptualizations considered perfectionism as unidimensional, studies that investigated the latent structure of different self-report measures have indicated the hegemonic presence of two-correlated dimensions.

The two dimensions of perfectionism received different nomenclatures over time, however, they are usually called Perfectionistic Strivings (PS) and Perfectionistic Concerns (PC) (Stoeber, 2018). These last nomenclatures seem to be more appropriate since they indicate that the two factors of perfectionism are two dimensions of the same construct, instead of two types of perfectionism, and for leaving the question of whether the dimensions are adaptive or maladaptive to empirical studies in this regard (Stoeber, 2016).

Regardless of the name they receive, the dimensions tend to encompass the same components. That is, PS embrace cognitions and behaviors of setting high personal demanding standards and pursuing them, whereas PC include cognitions and behaviors related to imperfections and their frightened consequences (Stoeber, 2016, 2018).

One of the main contributions of the two-dimensional model of perfectionism is the finding that each dimension presents distinct relationships with the psychopathological symptoms and conditions (Stoeber, 2016). Generally, PS are less related to psychopathology when compared to

---

Correspondence concerning this article should be addressed to: Luiz Fellipe Dias da Rocha, Universidade do Estado do Rio de Janeiro, Campus Francisco Negrão de Lima, 20550-900 Rio de Janeiro, Brasil. E-mail: luiz.rocha@uerj.br

PC (Limburg et al., 2016). A meta-analysis performed with 284 empirical studies – mostly cross sectional – revealed that both perfectionistic dimensions were positively related to psychological impairment indicators, it means, the higher the levels of PS and PC, the higher the level of psychopathological symptoms (e.g., depression, anxiety, and stress symptoms) (Limburg et al., 2016). Notwithstanding, PC's single effect values have shown to be very superior to PS's, after the control of these dimensions' overlap (β≤0.70 and β≤0.25, respectively).

Once perfectionism holds two dimensions with distinct relationships with psychopathological indicators, evaluating it accurately is essential to the advance of scientific knowledge regarding this field and to the development of an Evidence-Based Practice in Psychology. Measuring perfectionism (and its dimensions) precisely makes it possible to test theories and can be useful to suggest adding, subtracting, or prioritizing elements to an intervention method, making it more effective and efficient (see Suh et al., 2019). Among the available instruments to measure perfectionism, it lies the Clinical Perfectionism Questionnaire (CPQ) (Fairburn et al., 2003).

CPQ was developed aiming the evaluation of perfectionism from a cognitive-behavioral conceptualization, in which this construct is defined as the subordination of self-worth to the capacity of reaching high self-imposed standards (Shafran et al., 2002). Thereby, according to the perfectionistic cognitive processing, achieving high self-demanding standards guarantees personal value – despite temporarily, whereas not reaching them evokes thought about being a failure as a person (Fairburn et al., 2003; Shafran et al., 2002). Based on this, CPQ measures the tendency of a person to pursue self-demanding standards, as well as the cognitive and behavioral processes involved in this search and the subjective consequences from reaching standards or not.

CPQ is distinctive because it measures only the core characteristics of perfectionism, while many items from other scales do not evaluate integral elements of perfectionism, but assess related constructs, as beliefs about other people's standards and the perception that others pressure the individual to be perfect (Shafran et al., 2002). In addition, CPQ was developed according to a clinically based construct of perfectionism, that is, it captures the core psychopathological aspects of perfectionism, as morbid fear of failure, dichotomous thinking (i.e., operationalization of standards in the form of rules that are either met or not), and selective abstraction (i.e., paying more attention to negative perfectionism-relevant information than to positive information).

CPQ has 12 items, being two negatively keyed items (i.e., items 2 and 8). Negatively keyed items are those whose endorsement indicates a low level of the psychological variable being measured and, therefore, it must be reversed in the scoring process (Furr, 2011). In validation studies, the two negatively keyed items of CPQ demonstrated little impact on the global score through Corrected Item-Total Correlation (CITC), that is, the correlation of the individual item with the scale total omitting that item (Dickie et al., 2012; Howell et al., 2020; Moloodi et al., 2017; Rocha et al., 2022; Shu et al., 2019; Stoeber & Damian, 2014). It suggests that the negatively keyed items do not contribute significantly to the CPQ global score. Based on that, some authors (e.g., Dickie et al., 2012; Howell et al., 2020; Prior el al., 2018; Shu et al., 2019) have chosen to remove items 2 and 8.

Despite increasing the internal consistency of the instrument, excluding items exclusively based on CITC can provoke loss of potentially relevant information. High cohesion on a scale is important, but it is also important to consider that each item makes a different contribution to the scale. Therefore, internal consistency should not be the only indicator considered, but the variety in the item's contribution is also relevant. For example, item 2 of CPQ measures selective abstraction, a core characteristic in the conceptualization of clinical perfectionism. Thus, other resources can be used to increase the homogeneity of the scale without the need to discard items and risk losing potential data, such as rewording negatively keyed items into positively keyed items (Roszkowski & Soven, 2010).

Although some researchers include both positively and negatively keyed items attempting to reduce the acquiescence bias, doing so in an unbalanced manner might reduce the quality of the instrument (Furr, 2011; Roszkowski & Soven, 2010). A possible explanation for this occurrence is that the inclusion of some negatively keyed items in a mostly positive questionnaire seems to stimulate the tendency of a misinterpretation by the respondents, because they are being requested to shift gears in the cognitive processing a few times, what tends to create a response bias (Roszkowski & Soven, 2010). Furthermore, balancing the valence of the items do not achieve the goal of controlling the acquiescence bias, since respondents do not seem to process the information more deeply when there is the presence of negatively keyed items than when the instrument presents all items positively or negatively associated with the underlying concept (Menold, 2020).

Regarding latent structure, most studies with different samples found the presence of two factors, similar to PS and PC dimensions (Dickie et al., 2012; Egan et al., 2016; Moloodi et al., 2017; Rocha et al., 2022; Stoeber & Damian, 2014). It is, one factor comprises the items related to demanding standards, while the other embraces the items regarding failure and its consequences.

Dickie et al. (2012), through principal components analysis (PCA) and varimax rotation, found the two mentioned factors. Nevertheless, items 8 (*Over the past month, have you done just enough to get by? – negatively keyed items*) and 7 (*Over the past month, have you judged yourself on the basis of your ability to achieve high standards?*) were removed due to low item-total correlation and substantial crossloading (i.e., >.30), respectively. PCA from Stoeber and Damian (2014) found similar results, but with four out of the 12 items presenting substantial crossloading. These researchers performed an Exploratory Factor Analysis (EFA) and the results also revealed two factors, but the substantial crossloadings on items 7 and 8 persisted.

Egan et al. (2016), through EFA, carried out two studies: one with a non-clinical sample and the other with a sample of individuals with Eating Disorders. Solutions of two factors for CPQ were extracted in both studies, but items 1, 7 and 8 presented substantial crossloading. Moloodi et al. (2017), through Confirmatory Factor Analysis (CFA), have found similar results regarding CPQ psychometric properties, but without substantial crossloadings.

In the only study with a Brazilian sample (Rocha et al., 2022), EFA with promax rotation also indicated a two-factor solution, but item 8 was problematic regarding the factor analysis, because it did not present substantial factor loading in any of the two factors, being excluded in the subsequent analyzes. In addition, the unidimensionality of data was tested, and the results recommended that the instrument should not be treated as a predominantly unidimensional measure.

Despite the two-dimensional model, some studies have found a global factor for the instrument in addition to specific factors through bifactor approach (Howell et al., 2020; Prior et al., 2018). In a sample composed of patients with Eating Disorders, Prior et al. (2018) tested, through CFA, three models of CPQ: unidimensional, correlated two-factor and bifactor. Fit was revealed as poor for the first two models, and the bifactor model was not able to converge. After the exclusion of items 2 (*Over the past month, have you tended to focus on what you have achieved, rather than on what you have not achieved? – negatively keyed items*) and 8, due to low item-total correlation, a bifactor model presented a good fit to data: a global factor, with 10 items, and a specific factor (PS) with seven items.

Similarly, Howell et al. (2020) compared the unidimensional, correlated two-factor and bifactor models with scores from the 10-item version of CPQ (i.e., without items 2 and 8). The results also provided bigger support to the bifactor model, with a global and two specific factors. Additional analyses were performed in order to verify whether the instrument is predominantly uni-dimensional. Out of the four indicators generated to that end, two (i.e., Omega and H Coefficient) indicated a predominance of the global factor, while the other two (i.e., Percent uncontaminated correlations and explained common variance) did not confirm the unidimensionality of the instrument.

In summary, the factor structure of CPQ is not consensual. Most studies found two dimensions, as well as from other instruments that measure perfectionism. According to the results of these studies, PS and PC measured by the CPQ are two different factors, however related to one another – correlations between 0.23 and 0.48 (Dickie et al., 2012; Egan et al., 2016; Rocha et al., 2022 Stoeber & Damian, 2014). In another direction, two recent studies found that the bifactor model has indicated good fit to data based on CFA, which suggests the coexistence of specific and global factors (Howell et al., 2020; Prior et al., 2018). Thus, the two specific factors of the CPQ (i.e., PS and PC) could be combined into a general factor that holds all items.

Some studies have also investigated the validity of the CPQ regarding other variables, while comparing its scores with other instruments that measure maladjustment indicators, such as negative affects (Egan et al., 2016), depression, anxiety, and stress symptoms (Moloodi et al., 2017; Rocha et al., 2022; Shu et al., 2019) and eating disorders symptoms (Shu et al., 2019). In general, the PC dimension has shown a bigger relation with these indicators when compared to the PS dimension, as it occurred with other instruments that measure perfectionism (Limburg et al., 2016). The PC dimension of the CPQ also accounted for an additional variance in depression, anxiety and eating disorders symptoms, above and beyond another measure of perfectionism (Shu et al., 2019), showing good incremental validity.

Gilbert (2009, 2016), inspired in the affective neuroscience of evolution, describes two types of primary systems of emotional regulation (also called motivational systems) that operate in the humankind and might collaborate to understand the relationship between perfectionism factors and psychopathological indicators, they are: drive, seeking and acquisition focused system (or just drive system) and threat and self-protection focused system (or just threat system). The first one has as its function to motivate the pursuit of resources and rewards, guiding the development of desires and goals. The second one enables the individual to be alert and to quickly respond to threats, motivating security response in face of real or imaginary danger stimuli.

Even though the threat system has a primordially adaptive function, a lot of problems connected to mental health are related to it, especially if overdeveloped, sensitive, biased or confused (Gilbert, 2009). In general, the threat system is more related to psychopathology in comparison to the drive system due to the processes involved in its activation, such as neurohormonal mechanisms (e.g., cortisol response), negative emotions or negative emotional states (e.g., stress, fear, anxiety, and anger), biased cognitive processing (e.g., hypervigilance and worry) and often dysfunctional safety strategies (e.g., avoidance) (Gilbert, 2016; Irons & Beaumont, 2017). Nevertheless, the drive and threat systems interact in a way that the seek for resources might be connected to the attempt to fight against threats or to avoid them, designing a safety-seeking strategy to negative events (e.g., pursuing high standards to avoid inferiority feelings) (Gilbert, 2016).

Drawing a parallel with research on perfectionism, it is possible to observe similarities between the drive system and the PS dimension, once it could be seen as a mindset oriented to objectives that motivates people to pursue their high standards. On the other side, PC is similar to the threat system, since this dimension may be seen as a mindset that tries to avoid failures/imperfections and their negative consequences. Thereby, it would be possible to hypothesize that the PS dimension is related to the activation of the drive system, while the PC dimension is related to the activation of the threat system (Gilbert, 2009).

Considering that only one validation study with Brazilian samples was performed until the moment, that a version of CPQ with 12 positive items has not been tested yet and that the CPQ was able to explain variance in psychopathology, the present study aims at the verification of the psychometric properties of the Clinical Perfectionism Questionnaire (CPQ) in a Brazilian sample after rewording negative items into positive. The participation of the reworded items in the global score was investigated, in addition to test the latent structure through CFA and the correlation of CPQ with a psychopathological indicator.

Method

Participants

It was used a convenience sample of 168 Brazilian university students, aged between 18 and 56 years old (*M*=25.37, *SD*=7.40), being 76.8% women (*n*=129) and 22.6% men (*n*=38). One participant (0.6%) did not indicate sex. Regarding marital status, 81% declared to be single (*n*=136), 13.7% married (*n*=23), 4.7% divorced or separated (*n*=8) and 0.6% widow (*n*=1).

Instruments

The CPQ, cross culturally adapted by Rocha et al., 2022, was applied after rewording the two negatively keyed items into positive. In this study, it will be entitled CPQ+ (i.e., Clinical Perfectionism Questionnaire with 12 positive items). The wording of item 2 became '*Nos últimos 30 dias, você focou no que não conseguiu alcançar, ao invés de focar no que conseguiu alcançar*?'. Item 8 was rewritten as '*Nos últimos 30 dias, você deu o máximo de si em tudo o que fez*?'. Just like the original CPQ, the CPQ+ is composed of 12 items that examine the frequency of some behaviors and cognitions related to perfectionism in the last 30 days. The respondent must indicate the answer through a Likert-type scale from 1 (Not at all) to 4 (All of the time).

The Depression, Anxiety and Stress Scales – Short Form (DASS-21) by Lovibond and Lovibond (1995), adapted to Brazilian population by Vignola and Tucci (2014), were also applied. This self-report instrument is composed of 21 items that measure depression, anxiety and stress symptoms. The respondents must indicate how much each item was related to their reality in the last week. Answers are given in Likert-type scale, from "Did not apply at all" (0) to "Applied a lot or most of the time" (3). Although the presence of specific factors (i.e., depression, anxiety, and stress), the measure is predominantly unidimensional (i.e., negative affectivity) (Rocha et al., 2021).

Proceedings

After approval of the project by the Ethics in Research Committee of the institution to which this study is attached, an online questionnaire was created through Google Forms, containing the Written Informed Consent Form (WICF) and the research instruments. The invitation to participate in the project was made through social media ads (e.g., Facebook) and e-mails to university professors, requesting the forwarding of the form link to their students. The resource of mandatory responses was used for all items, which prevented the form from being returned with missing values. Respondents took an average of 12 minutes and 44 seconds to complete the form.

Data analysis

The obtained data were typed in the SPSS (version 23) and, firstly, a descriptive statistical analysis was performed. In the next stage, the CITC of each item of the CPQ+ was investigated, it means, the correlation between each item and a sum of the others (excluding that item). The CITC corresponds to an indicator applied to verify whether each item contributes substantially to the global score of the scale, being recommended values above .30 (Streiner et al., 2015).

In the Analysis of Moment Structures (Arbuckle, 2017) software, through CFA, the unidimensional, correlated two-factor and bifactor models were tested with the Maximum Likelihood estimate method, which is shown to be robust, even in the presence of a non-normal

distribution of data (Marôco, 2014). In order to evaluate the current fit of the examined models, the same indexes reported by other studies were taken into consideration (i.e., Howell et al., 2020; Moloodi et al., 2017; Prior et al., 2018; Shu et al., 2019).

The chi square (χ^2) evaluates the discrepancy magnitude between the population covariance matrix and the sample covariance matrix. It is a conservative estimate of the model fit when the size of the sample is smaller than 200 (Byrne, 2016). In this case, the χ^2/df ratio must be applied and the results inferior to 2 or 3 are considered good ones. Comparative Fit Index (CFI) compares the evaluated model fit with the basal model and values bigger than .90 indicate a good fit (Bentler, 1990; Marôco, 2014). Goodness-of-Fit Index (GFI) and Adjusted Goodness-of-Fit Index (AGFI) are adjustment measures between the hypothetical model and the observed covariance matrix, and values bigger than .90 are considered good and bigger than .95 are very good. Incremental Fit Index (IFI) is a relative adjustment index that compares the χ^2 to the tested model with the so-called null model, and values bigger than .90 are considered good and bigger than .95 are very good. Root Mean Square Error of Approximation (RMSEA) measures the discrepancy through degrees of freedom between sample and population estimates, and values smaller than .05 are considered very good (Hair et al., 2018).

Furthermore, the Akaike Information Criterion (AIC) was added, which is an index based on the χ^2 statistics and penalized the model according to its complexity. The best fit will present the lowest values (Arbuckle, 2017). Internal consistency was calculated through Composite Reliability (CR) and Cronbach's alpha (α).

In the classical statistics area, in order to seek validity in relation to other variables, zero-order correlations of CPQ+ scores with DASS-21 global score were performed, once these instruments measure related variables. The literature on the relation between perfectionism and negative affectivity is extensive, permitting comparison. Moreover, no other validated instruments to measure perfectionism in Brazilians were found. The Meng and colleagues (1992) method of comparing sets of non-independent correlations was used to compare the correlations between CPQ+ factors and DASS-21 global score.

Results

Analysis of multivariate distribution of scores revealed its non-normality: Mardia's Coefficient was 3.21 (normalized=1.14). The observation of Mahalanobis distance did not identify the presence of multivariate outliers in data. Furthermore, at the univariate distribution, the asymmetry values were <±0.8 and kurtosis <±1.1, which is not an extreme normality violation (Tabachnick & Fidell, 2018).

The results of the descriptive analyzes are shown in Table 1, as well as the CITC of the CPQ. As can be seen, items presented coefficients between 0.29 (item 08) and 0.62 (item 5). These results suggest that all items substantially contribute to the global score of CPQ+ (i.e., CITC≥.30), although item 8 has been slightly below expected.

CFA of unidimensional and correlated two-factor models of CPQ+ with 12 items revealed inappropriate general fit to empirical data (Table 2). Due to the simultaneous saturation in both factors of CPQ+, Modification Indexes suggested the elimination of item 8. It was also indicated the correlation between errors of items 3 and 10. After these re-specifications, the correlated two-factor model of CPQ+ with 11 items (Figure 1) showed a good general fit (Marôco, 2014) to empirical data (Table 2).

Table 1

Descriptive analyzes of the variables and CITC

	CITC	Mean	SD	95% CI
Item 1. Nos últimos 30 dias, você se pressionou muito para atingir seus objetivos?	.49	2.70	.80	2.57, 2.82
Item 2. Nos últimos 30 dias, você focou no que não conseguiu alcançar, ao invés de focar no que conseguiu alcançar?	.40	2.53	.86	2.40, 2.66
Item 3. Nos últimos 30 dias, alguém lhe disse que seus padrões de exigência são muito elevados?	.55	2.20	1.05	2.04, 2.36
Item 4. Nos últimos 30 dias, você se sentiu um fracasso como pessoa por não ter conseguido atingir seus objetivos?	.51	2.70	.94	2.56, 2.85
Item 5. Nos últimos 30 dias, você teve medo da possibilidade de não alcançar seus padrões de exigência?	.62	3.08	.89	2.94, 3.21
Item 6. Nos últimos 30 dias, você aumentou seus padrões de exigência por achá-los fáceis demais?	.49	1.96	.88	1.82, 2.09
Item 7. Nos últimos 30 dias, você avaliou seu valor com base em sua habilidade de atingir elevados padrões de exigência?	.61	2.52	.91	2.38, 2.66
Item 8. Nos últimos 30 dias, você deu o máximo de si em tudo o que fez?	.29	2.60	.81	2.48, 2.72
Item 9. Nos últimos 30 dias, você checou repetidamente o quão bom você é em atingir seus padrões de exigência (por exemplo, comparando seu desempenho ao dos outros)?	.60	2.62	.96	2.47, 2.77
Item 10. Nos últimos 30 dias, você acha que as pessoas te consideraram perfeccionista?	.51	2.11	.89	1.97, 2.24
Item 11. Nos últimos 30 dias, você continuou tentando alcançar seus padrões de exigência, mesmo que para isso tenha que ter aberto mão de algumas coisas?	.46	2.40	.86	2.27, 2.54
Item 12. Nos últimos 30 dias, você evitou pôr sua capacidade à prova por medo de falhar?	.32	2.37	.94	2.23, 2.51
CPQ+ total	-	29.78	6.39	28.80, 30.75
CPQ+ PS	-	16.50	4.27	15.85, 17.15
CPQ+ PC	-	10.68	2.73	10.26, 11.09
DASS-21 total	-	24.51	14.12	22.36, 26.66

Note. CITC=corrected item-total correlation. *SD*=Standard derivation. 95% CI=95% confidence interval. CPQ+ total=sum of the 12 items. CPQ+ PS=sum of items 1, 3, 6, 7, 9, 10 and 11. CPQ+ PC=sum of items 2, 4, 5 and 12. DASS-21 total=sum of the 21 items.

Table 2

General fit indexes of CFA of different CPQ models previously and currently tested

Study	Model	χ^2/df	CFI	GFI	AGFI	IFI	RMSEA (90%CI)	AIC
Current study	12 items							
	Unidimensional	4.29	.70	.78	.69	.71	.14 (.12. .16)	279
	Correlated two-factor	3.3	.80	.84	.76	.80	.12 (.10, .14)	225
	Bifactor*	1.5	.97	.95	.90	.97	.05 (.02, .08)	131
	11 items							
	Correlated two-factor	1.8	.94	.92	.88	.94	.07 (.04, .09)	123
Howell et al. (2020)	10 items							
	Unidimensional	7.36	.88				.127 (.112, .141)	
	Correlated two-factor	5.44	.91				.108 (.093, .123)	
	Bifactor	2.86	.97				.068 (.050, .087)	
Shu et al. (2019)	12 items							
	Unidimensional	5.3	.70	.81	.72		.126 (.112, .141)	331
	Correlated two-factor	3.5	.83	.89	.84		.097 (.082, .112)	235
	10 items							
	Correlated two-factor	2.4	.94	.94	.91		.072 (.052, .093)	123
Prior et al. (2018)	12 items							
	Unidimensional	6.36	.85				.153 (.138, .169)	
	Correlated two-factor	4.76	.90				.128 (.113, .144)	
	10 items							
	Unidimensional	6.25	.91				.151 (.132, .171)	
	Correlated two-factor	3.8	.95				.111 (.091, .131)	
	Bifactor	2.73	.97				.090 (.067, .113)	
Moloodi et al. (2017)	11 items							
	2 F oblíquos	1.8	.91	.97	.95	.92	.04 (.02, .06)	

Note. *Indexes before the exclusion of non-significant items (*p*>.05). Eliminating these non-significant loadings, the bifactor model fit has become unfeasible.
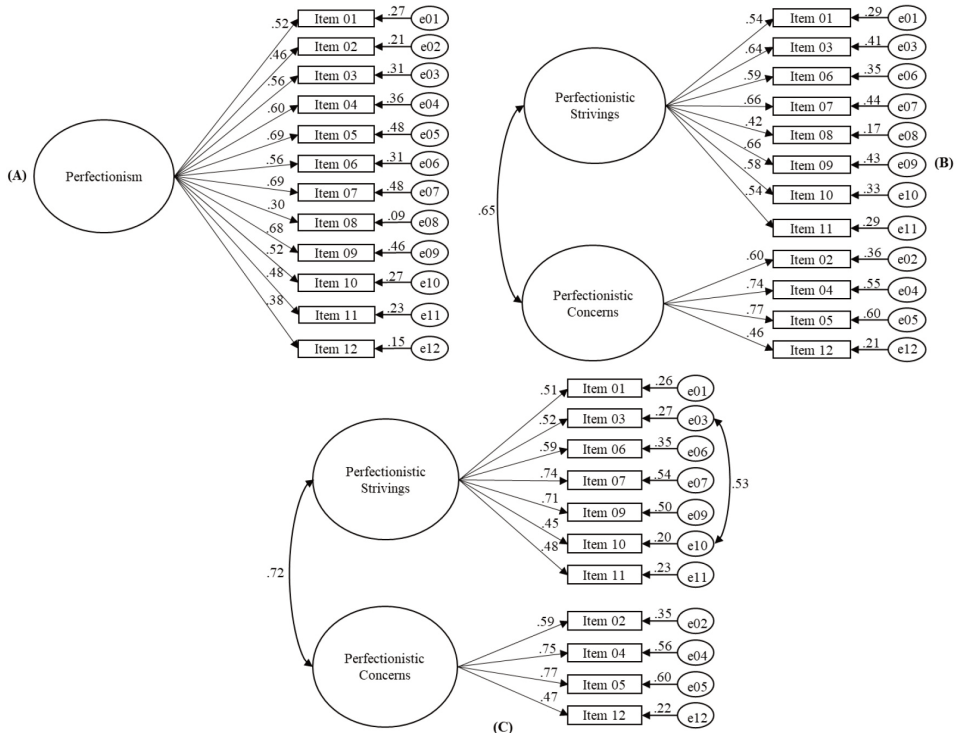
Figure 1. Path diagrams of the different CPQ+ factorial models analyzed through CFA: (A) Unidimensional with 12 items, (B) Correlated two-factor with 12 items, (C) re-specified correlated two-factor with 11 items

Nevertheless, the observation of the correlated two-factor model of CPQ+ identified that items 10 and 11, from PS, and item 12, from PC, remained with factor weights (λ) <0.50 (Figure 1). This reflected an insufficient individual reliability of these items. Hair et al. (2018) considered that $λ^2 ≥ 0.25$ is an appropriate individual reliability indicator and would represent a minimal explanation of the item's variance through latent factor.

The convergent validity in the Structural Equation Modeling (SEM) field is represented by strong saturation (≥0.50) of items on its factors and, consequently, it is expected that the average variance extracted (AVE) of each factor is ≥0.50 (Hair et al., 2018). In the current research, the AVE calculated for the PS and PC factors were 0.34 and 0.43, respectively, that represent insufficient convergent validities. On the other side, CR were 0.78 and 0.75, and αs were 0.79 and 0.74, respectively, which is an adequate reliability for both factors.

The comparison of the correlation square of factors with AVE values of each factor is a discriminative validity indicative in the SEM field. It was obtained a correlation of 0.73 between CPQ+ factors, that squared, $R^2 = 0.53$. As the values calculated for the AVE of factors were less than its correlation square, it was not sufficiently investigated the discriminative validity between the two dimensions of the instrument (Hair et al., 2018; Marôco, 2014).

Considering these restrictive preliminary results, the bifactor model testing was proceeded to investigate the factor structure of the CPQ+. Although item 8 has presented difficulties in previous tests and was excluded from subsequent analysis, it was included again. The bifactor model was tested with 12 items. Although, apparently, indexes have shown a very good general fit to data, the observation of item's estimate revealed that a great part was not significant ($p > 0.05$). Eliminating these non-significant loadings, the bifactor model adjustment has become unfeasible.

Moreover, the AIC value of the correlated two-factor model was inferior to the one of the bifactor model before the non-significant loadings, indicating the first's best fit.

In the classical statistics area, zero-order correlations showed a bigger relationship between PC and DASS-21 total score (r=.574, p<.01), in comparison to PS and DASS-21 total score (r=.448, p<.01). As can be seen, the numerical difference between the two correlations was equal to Δr=.119. The 95% confidence intervals associated with the difference in the correlation was estimated at r=.007 and r=.320. The null hypothesis of equal zero-order correlations was rejected based on one-tailed p-value testing (z=1.877, p<.05).

Discussion

The present study aimed at the verification of the psychometric properties of the Clinical Perfectionism Questionnaire (CPQ) in a Brazilian sample after rewording negatively keyed items into positively keyed items. For that, the participation of the reworded items in the global score was investigated through the CITC calculation, in addition to the analysis of the latent structure of this new version of the instrument through CFA and the correlation of its scores with psychopathological indicators.

The CITC values of the reworded items (i.e., items 2 and 8) were better than those found by other studies that used items in its original form (i.e., negatively keyed items) (Dickie et al., 2012; Moloodi et al., 2017; Rocha et al., 2022; Shu et al., 2019; Stoeber & Damian, 2014), although the coefficient of item 8 has been slightly beneath .30. As suggested by Roszkowski and Soven (2010), using items written only in the same direction (i.e., only positively keyed items) might make the instrument more homogeneous, tending to improve its internal consistency. Moreover, higher CITC values might demonstrate a better comprehension of items by the respondents (Cohen et al., 2014), which is expected to reduce a possible effect of response bias.

It is true that some researchers include negatively keyed items in scales to guard against problems, such as participants' acquiescing. However, empirical evidence indicates that negatively keyed items may be creating validity and reliability problems, in addition to not achieve the goal of controlling the acquiescence bias (Menold, 2020).

Although some authors (e.g., Dickie et al., 2012; Howell, Anderson, Egan, & McEvoy, 2020; Prior el al., 2018; Shu et al., 2019) have removed items 2 and 8 from CPQ based on the low CITC, excluding items based on that coefficient can provoke loss of potentially relevant information. For example, item 2 measures selective abstraction, a core characteristic in the conceptualization of clinical perfectionism. So, rewording items is a better option than excluding them.

Regarding the CPQ+ factor structure, the current results revealed that the test of the correlated two-factor model with 11 items presented a good fit to data, according to Marôco's classification (Marôco, 2014), overcoming the fit of other models tested through the AIC calculation. After the removal of item 8, with low CITC, Moloodi et al. (2017) also found a good fit to this same model with 11 items through CFA. Similarly, Shu et al. (2019) eliminated items 2 and 8 and found the same result, but with 10 items. The results obtained in the present study through CFA also corroborated the two-dimensional model of perfectionism indicated by previous studies that performed EFA (Dickie et al., 2012; Egan et al., 2016; Rocha et al., 2022; Stoeber & Damian, 2014), and by other instruments that measure perfectionism (see Stoeber, 2016, 2018).

The path representing the covariance (correlation) among error terms of items 3 and 10 was inserted according to modification indexes. Within-construct error variance covariance (i.e., covariance among error terms of items indicating the same construct) might suggest the existence of another factor, not directly measured by the instrument, explaining the relationship between

the items (Hair et al., 2018). In the case of items 3 and 10, both make references to attitudes that the respondents assign to others regarding themselves (i.e., meta-perception). Meta-perception is commonly used as a methodological approach to contextualize personality assessment, once it aggregates relevant information about how the respondent perceives the manifestations of the personality trait in interpersonal contexts – rather than taking a purely self-evaluative approach (see Woods, 2004).

The correlated two-factor model, in the present study, still did not present enough convergent and discriminative validity. The convergence of items might be expressed by the amount of common variance shared by items, that ideally should overcome the amount of variance of error. However, the measures of internal consistency, CC and α, that also represent convergence, presented proper indexes. The discriminative validity in CFA is the degree to which a factor is different from the others (Hair et al., 2018).

These limitations, found in the correlated two-factor model, introduced the testing of the bifactor model, that, although has reached a very good general fit to data, became unfeasible after the elimination of the non-significant loadings. Even before this elimination in the bifactor model, the correlated two-factor model presented a superior fit when the AIC value of both were compared.

It is important to highlight that the correlation value between factors in this study was substantially higher to those found in other studies (Dickie et al., 2012; Egan et al., 2016; Rocha et al., 2022; Shu et al., 2019; Stoeber & Damian, 2014), which might be caused due to the characteristics of the sample employed, and that this might have influenced the indicators of convergent and discriminative validity.

It is a fact that the CPQ was designed according to a unidimensional conceptualization as from the critics of the multidimensional approach to perfectionism (see Shafran et al., 2002). Nevertheless, empirical studies have shown that the factor structure of this measure has correlated two-factor, being one regarding the PS dimension and the other concerning the PC dimension, reinforcing the two-dimensional model of perfectionism (Stoeber, 2016, 2018). Rocha et al. (2022) tested, based on the conceptualization of Clinical Perfectionism, if the CPQ scores could be treated as predominantly unidimensional, but the results did not support this hypothesis.

An important contribution from the two-dimensional model of perfectionism is the findings that the dimensions have a distinct relation with psychopathological indicators (Limburg et al., 2016). In this field, CPQ holds important relevance, once it has accounted for an additional variance in psychopathology above and beyond other measures of perfectionism (Shu et al., 2019), maybe because their authors have taken into consideration the clinically relevant components of perfectionism during the construction of the instrument (Shafran et al., 2002).

In this study, the PC dimension was more related to negative affectivity measured by DASS-21 in comparison to the PS dimension, indicating good validity in relation to other variables. This result was already expected, based on similar studies (Egan et al., 2016; Limburg et al., 2016; Moloodi et al., 2017; Shu et al., 2019) and on the primary systems of emotional regulation model proposed by Gilbert (2009).

High levels of PS revealed a tendency of the individual to pursue high demanding standards through the activation of the drive system, which tends to generate more pleasant than unpleasant emotions, as feeling energized to achieve and satisfied when achieved (Gilbert, 2009; Irons & Beaumont, 2017). On the other hand, high levels of PC make the individuals interpret errors as threats to the self-worth and this type of information processing activates the threat system, which generates negative affectivity symptoms (Gilbert, 2009; Irons & Beaumont, 2017). Thus, although both dimensions can be correlated, the affectivity outcomes from each one are not the same.

Nevertheless, the drive and threat systems interact in a way that the pursuit of resources might be connected to the attempt to fight against threats or to avoid them, designing a safety strategy to negative events (Gilbert, 2016). In practical terms, interventions that seek to regulate

perfectionism, in order to reduce negative affectivity, should focus on reducing PC dimension levels, not losing sight of the use of the PS dimension as safety strategies, that is, behaviors that are performed to avoid a frightened result or minimize its adverse consequences.

One of the problems with safety strategies is that they prevent the individual from an accurate perception of threats and from developing more functional strategies to deal with them (see Gray et al.2019). In the case of PC, the failure tends to awaken negative thoughts about oneself, as believing to be a failure as a person, that, on its turn, will lead to the experience of negative emotions. Naturally, the person will be afraid of failing and will try to avoid the failure, actively (e.g., trying to reach perfection) or passively (e.g., procrastinating or avoiding testing one's capacity) (Shafran et al., 2002).

Despite the relevance of the results presented, some limitations of the present study must be considered. It is evident that the small, non-probabilistic, and by convenience sample employed does not represent the Brazilian population, hindering the extrapolation of results. Moreover, studies with clinical Brazilian samples will also be important to the classification of the factor structure of CPQ+ and its relationship with psychopathological indicators. The relations between CPQ scores and positive indicators, such as subjective wellbeing, lack investigation and were neither encompassed here. At last, experimental studies with native samples will be able to indicate whether CPQ scores are sensitive to the effects of treatment and which interventions are more efficient at the reduction of perfectionism and its outcomes within this context. The continuation of investigations will be able to project light to these questions.

Conclusions

Generally, in the present study, rewording the negatively keyed items seemed to be a good strategy to increase the homogeneity of the instrument without deleting items. In the testing of multiple models of CPQ+, the best fit to data remained with the correlated two-factor model with 11 items. This result is coherent with the two-dimensional model of perfectionism and with previous empirical tests. Furthermore, the PC factor has shown to be more related to psychopathological symptoms in comparison to the PS factor, which is also coherent with the theoretical model proposed by Gilbert (2009), and with previous studies employing CPQ and other instruments to measure perfectionism.

Funding

Author contributions

Conceptualization: LFDR, EMOF; Methodology: LFDR, JAEH; Project administration: LFDR; Software: LFDR, JAEH; Validation: LFDR, JAEH; Visualization: LFDR, JAEH, EMOF; Formal analysis: LFDR, JAEH; Investigation: LFDR; Data curation: LFDR; Writing – Original draft preparation: LFDR; Writing – Review and editing: LFDR, JAEH, EMOF; Supervision: JAEH, EMOF.

All authors read and approved the final manuscript.

References

Arbuckle, J. L. (2017). *IBM® SPSS® AMOSTM 25.0: User's guide*. Amos Development Corporation.

Bentler, P. M. (1990). Comparative fit indexes in structural models. *Psychological Bulletin, 107*(2), 238-246. https://doi.org/10.1037/0033-2909.107.2.238

Byrne, B. (2016). *Structural equation modeling with AMOS: Basic concepts, applications, and programming.* Routledge. https://doi.org/10.4324/9781315757421

Cohen, R. J., Swerdlik, M. E., & Sturman, E. D. (2014). *Testagem e avaliação psicológica: Introdução a testes e medidas*. AMGH.

Dickie, L., Surgenor, L. J., Wilson, M., & McDowall, J. (2012). The structure and reliability of the Clinical Perfectionism Questionnaire. *Personality and Individual Differences, 52*(8), 865-869. https://doi.org/10.1016/j.paid.2012.02.003

Egan, S. J., Shafran, R., Lee, M., Fairburn, C. G., Cooper, Z., Doll, H. A., Palmer, R. L., & Watson, H. J. (2016). The reliability and validity of the clinical perfectionism questionnaire in eating disorder and community samples. *Behavioural and Cognitive Psychotherapy, 44*(1), 79-91. https://doi.org/10.1017/S1352465814000629

Fairburn, C. G., Cooper, Z., & Shafran, R. (2003). *The Clinical Perfectionism Questionnaire*. Unpublished manuscript. Department of Psychiatry, University of Oxford, UK.

Furr, M. (2011). *Scale construction and psychometrics for social and personality psychology*. Sage publications.

Gilbert, P. (2009). *The compassionate mind: A new approach to life's challenges*. Robinson.

Gilbert, P. (2016). *Human nature and suffering*. Routledge.

Gray, E., Beierl, E. T., & Clark, D. M. (2019). Sub-types of safety behaviours and their effects on social anxiety disorder. *Plos One, 14*(10), 1-18. https://doi.org/10.1371/journal.pone.0223165

Hair, J. F., Black, W. C., Babin, B. J., & Anderson, R. E. (2018). *Multivariate data analysis* (8th ed.). Cengage Learning EMEA.

Howell, J., Anderson, R., Egan, S., & McEvoy, P. (2020). One factor? Two factor? Bi-factor? A psychometric evaluation of the Frost Multidimensional Scale and the Clinical Perfectionism Questionnaire. *Cognitive Behaviour Therapy, 49*(6), 518-530. https://doi.org/10.1080/16506073.2020.1790645

Irons, C., & Beaumont, E. (2017). *The compassionate mind workbook: A step-by-step guide to developing your compassionate self*. Robinson.

Limburg, K., Watson, H. J., Hagger, M. S., & Egan, S. J. (2016). The relationship between perfectionism and psychopathology: A meta-analysis. *Journal of Clinical Psychology, 73*(10), 1301-1326. https://doi.org/10.1002/jclp.22435

Lovibond, P. F., & Lovibond, S. H. (1995). The structure of negative emotional states: Comparison of the Depression Anxiety Stress Scales (DASS) with the Beck Depression and Anxiety Inventories. *Behaviour Research and Therapy, 33*(3), 335-343. https://doi.org/10.1016/0005-7967(94)00075-u

Marôco, J. (2014). *Análise de equações estruturais: Fundamentos teóricos, software & aplicações* (2ª ed.). ReportNumber.

Meng, X. L., Rosenthal, R., & Rubin, D. B. (1992). Comparing correlated correlation coefficients. *Psychological Bulletin, 111*(1), 172-175. https://doi.org/10.1037/0033-2909.111.1.172

Menold, N. (2020). How do reverse-keyed items in inventories affect measurement quality and information processing?. *Field Methods, 32*(2), 140-158. https://doi.org/10.1177/1525822X19890827

Moloodi, R., Pourshahbaz, A., Mohammadkhani, P., Fata, L., & Ghaderi, A. (2017). Psychometric properties of the Persian version of clinical perfectionism questionnaire: Findings from a clinical and non-clinical sample in Iran. *Personality and Individual Differences, 119*, 141-146. https://doi.org/10.1016/j.paid.2017.07.003

Prior, K. L., Erceg-Hurn, D. M., Raykos, B. C., Egan, S. J., Byrne, S., & McEvoy, P. M. (2018). Validation of the clinical perfectionism questionnaire in an eating disorder sample: A bifactor approach. *International Journal of Eating Disorders, 51*(10), 1176-1184. https://doi.org/10.1002/eat.22892

Rocha, L. F. D., Hernandez, J. A. E., & Falcone, E. M. O. (2021). Latent structure evidence of the Depression, Anxiety and Stress Scales – Short form. *Estudos de Psicologia (Campinas), 38*, 1-13. https://doi.org/10.1590/1982-0275202138e190103

Rocha, L. F. D., Hernandez, J. A. E., Falcone, E. M., & Peluso, M. L. (2022). Cross-cultural adaptation of the Clinical Perfectionism Questionnaire (CPQ) for Brazilians. *PsicoUSF, 27*(1), 73-85. https://doi.org/10.1590/1413-82712022270106

Roszkowski, M. J., & Soven, M. (2010). Shifting gears: Consequences of including two negatively worded items in the middle of a positively worded questionnaire. *Assessment & Evaluation in Higher Education, 35*(1), 113-130. https://doi.org/10.1080/02602930802618344

Shafran, R., Cooper, Z., & Fairburn, C. G. (2002). Clinical perfectionism: A cognitive-behavioural analysis. *Behaviour Research and Therapy, 40*(7), 773-791. https://doi.org/10.1016/S0005-7967(01)00059-6

Shu, C. Y., O'Brien, A., Watson, H. J., Anderson, R. A., Wade, T. D., Kane, R. T., Lampard, A., & Egan, S. J. (2019). Structure and validity of the Clinical Perfectionism Questionnaire in female adolescents. *Behavioural and Cognitive Psychotherapy, 48*(3), 268-279. https://doi.org/10.1017/s1352465819000729

Stoeber, J. (2016). Perfectionism. In V. Zeigler-Hill & T. K. Shackelford (Eds.), *Encyclopedia of personality and individual differences*. Springer. https://doi.org/10.1007/978-3-319-28099-8_2027-1

Stoeber, J. (2018). The psychology of perfectionism: An introduction. In J. Stoeber (Ed.), *The psychology of perfectionism: Theory, research, applications* (pp. 3-16). Routledge.

Stoeber, J., & Damian, L. E. (2014). The Clinical Perfectionism Questionnaire: Further evidence for two factors capturing perfectionistic strivings and concerns. *Personality and Individual Differences, 61-62*, 38-42. https://doi.org/10.1016/j.paid.2014.01.003

Streiner, D. L, Norman, G. L., & Cairney, J. (2015). *Health measurement scales: A practical guide to their development and use* (5th ed.). Oxford University Press.

Suh, H., Sohn, H., Kim, T., & Lee, D. G. (2019). A review and meta-analysis of perfectionism interventions: Comparing face-to-face with online modalities. *Journal of Counseling Psychology, 66*(4), 473-486. https://doi.org/10.1037/cou0000355

Tabachnick, B. G., & Fidell, L. S. (2018). *Using Multivariate Statistics* (7th ed.). Pearson.

Vignola, R. C. B., & Tucci, A. M. (2014). Adaptation and validation of the Depression, Anxiety and Stress Scale (DASS) to Brazilian Portuguese. *Journal of Affective Disorders, 155*, 104-109.

Woods, S. A. (2004). *Contextualising personality assessment: Using meta-perceptions of the Big Five to improve the prediction of work performance* [Doctoral dissertation, University of Surrey]. https://openresearch.surrey.ac.uk/esploro/outputs/doctoral/Contextualising-Personality-Assessment-Using-Meta-perceptions-of/99511109202346#file-0

Propriedades psicométricas da versão brasileira do Clinical Perfectionism Questionnaire após reformulação de itens com semântica negativa

Resumo: O presente estudo tem como objetivo verificar as propriedades psicométricas da versão brasileira do Questionário de Perfeccionismo Clínico após a reformulação de itens negativos em

positivos (CPQ+). Foi utilizada uma amostra de conveniência de 168 universitários brasileiros (*M*=25.37, *DP*=7.40), sendo 76.8% mulheres. Após reformular os itens negativos para positivos, o CPQ+ foi aplicado. A Depression, Anxiety and Stress Scales – Short Form (DASS-21) também foi aplicada. Os valores de Correlação Item-Total Corrigido dos itens do CPQ+ foram limítrofes ou superiores ao ponto de corte. Após a exclusão do item 8, devido ao crossloading, e a inclusão da correlação entre os erros dos itens 3 e 10, o modelo de dois fatores correlacionados foi o de melhor ajuste para o CPQ+ através da Análise Fatorial Confirmatória. A dimensão Preocupações Perfeccionistas do CPQ+ esteve mais relacionada ao escore global da DASS-21 em comparação à dimensão Esforço Perfeccionista, indicando boa validade em relação a outras variáveis. Em síntese, os resultados encontrados recomendam a utilização do CPQ+ apresentado neste estudo, ou seja, uma versão com 11 itens positivos.