

Forms of Self-Criticising/Attacking & Self-Reassuring Scale: psychometric properties and normative study

Rita Baião¹, Paul Gilbert², Kirsten McEwan², & Sérgio A. Carvalho³

¹ Psychology Research Centre (CIPsi), University of Minho

² School of Sciences, University of Derby

³ Cognitive and Behavioural Research Centre (CINEICC), University of Coimbra

Abstract

The Forms of Self-Criticising/Attacking & Self-Reassuring Scale (FSCRS, Gilbert, Clarke, Hempel, Miles & Irons, 2004) is a self-report instrument that measures self-criticism and self-reassurance. It has shown good reliability and has been used in several different studies and in a range of different populations. The aim of the present study is to explore its psychometric proprieties in a large clinical and nonclinical sample, in order to provide reliability and, for the first time, normative data. Differences in population scores will also be addressed.

Method: Data was collated from 12 different studies, resulting in 887 nonclinical participants and 167 mixed diagnosis patients who completed the FSCRS.

Results: A confirmatory factor analysis shows that both in non-clinical and clinical samples, the three-factor model of FSCRS is a well-adjusted measure for assessing the two forms of self-criticism and a form of self-reassurance. Normative data for the scale is presented. Comparing the two populations, the nonclinical was more self-reassuring and less self-critical than the clinical. Comparing genders, in the nonclinical population men were more self-reassuring and less self-critical than women. No significant gender differences were found in the clinical population.

Conclusions: Taken together, results corroborate previous findings about the link between self-criticism and clinical population, which stresses the need to assess it. Results also confirm that FSCRS is a robust and reliable instrument, which now can aid clinicians and researchers to have a better understanding of the results, taking into account the norms presented.

Keywords: self-criticism, FSCRS, confirmatory factor analysis, normative study

Practitioner Points

Practical implications:

- The normative study of the FSCRS facilitates a better understanding of clinical and research results;
- The paper accounts for large clinical and nonclinical populations, which contribute to robust findings;

Cautions:

- Cultural and age differences should be carefully addressed;
- Generalizations to different psychopathologies deserve attention, as the clinical population considered here derived mainly from depressed participants.

Introduction

Self-criticism is one of the most pervasive features of psychopathology (Gilbert & Irons, 2005; Zuroff, Santor, & Mongrain, 2005). It is highly associated with shame (Gilbert, McEwan, Gibbons, Chotai, Duarte, et al., 2012) which is another prominent feature of psychopathology. Its pathogenic qualities may derive from the strength of negative emotions related to it, especially (self-directed) anger, disgust and contempt (Whelton & Greenberg, 2005), and their link to emotional memories (Gilbert, 2010). Self-criticism has a negative impact on psychological interventions. For example, Rector, Bagby, Segal, Joffe, and Levitt (2000), found that self-critical patients were more likely to have a poor response to Cognitive Behavioural Therapy (CBT). They also found that successful treatment responses were associated with significant reductions in self-criticism. In the same way, in a study consisting of 84 outpatients with social phobia, Cox, Walker, Enns, & Karpinski (2002) found that changes in levels of self-criticism were significantly associated with positive responses to the social phobia treatment.

Self-criticism is associated with activity in lateral prefrontal cortex (PFC) regions and dorsal anterior cingulate, linking self-critical thinking to error processing and resolution, and also behavioural inhibition (Longe, Maratos, Gilbert, Evans, Volker, et al., 2010). Longe, et al. (2010) also found that dorsolateral PFC activity was positively correlated with high levels of self-criticism, suggesting again greater error

processing and behavioural inhibition in those individuals. In contrast, the ability to be self-reassuring and self-compassionate stimulate different brain systems (Longe, et al 2010) and is negatively linked to psychopathology (Gilbert, Clarke, Hempel, Miles, & Irons, 2004; Neff, 2003).

The *Forms of Self-Criticising/Attacking & Self-Reassuring Scale* (FSCRS, Gilbert, et al., 2004) was developed to explore different ways people treat themselves when things go wrong - in particular measuring tendencies to be self-critical and/or self-reassuring when perceiving setbacks/failures. Items derived from clinical practice, based on thoughts depressed patients presented about their own self-criticism and ability to self-reassure. Factor analysis suggested one factor of self-reassurance, and two different factors of self-criticism (one focused on feeling inadequate, and another one related to a more self-hating and contemptuous feelings of self).

The original study of FSCRS' psychometric properties was conducted on a sample of 246 female undergraduate students (Gilbert, et al., 2004). The two self-critical subscales are considered to relate to psychopathology in different ways, with the self-hating dimension representing a more pathological domain associated with self-harm and more borderline phenomenology (Gilbert, et al., 2004, Gilbert, McEwan, Irons, Bhundia, Christie, et al., 2010). The two subscales have also shown different distribution of responses, with only hated-self showing a floor effect in nonclinical samples (Gilbert et al., 2012).

Nonetheless, these two subscales have been shown to be strongly correlated with each other (r from .68 to .80) (e.g. Gilbert, et al., 2004, 2010; Irons, Gilbert, Baldwin, Baccus & Palmer, 2006; Richter, Gilbert, & McEwan, 2009). Therefore, some studies have combined the two subscales into one factor of self-criticism (e.g. Gilbert, Baldwin, Irons, Baccus, & Palmer, 2006), particularly for the ease of investigating the mediator/moderator effects of self-criticism on psychopathology variables (Richter et al., 2009). Recently, a study on the FSCRS as part of an online survey (N = 1,570) confirmed the three-factor structure of the scale, and the two different types of self-criticism (Kupeli, Chilcot, Schmidt, Campbell, & Troop, 2013). Also, a different study, based on the Portuguese version of the FSCRS, found good psychometric characteristics, and the three factors discriminating between the clinical and nonclinical samples (Castilho, Pinto-Gouveia & Duarte, 2013). Therefore, the 3-factor structure seems to replicate well in different samples.

In fact, the FSCRS has been used in a range of different studies, in which self-criticism has been linked to depression and anxiety (e.g. Gilbert et al., 2004), self-harm (Gilbert, et al., 2010), negative future thinking (Goodall, Gilbert, & McEwan, submitted), early memories of threat and submissiveness (Richter, et al., 2009), fears of compassion (Longe et al., 2010; Rockliff, Gilbert, McEwan, Lightman, & Glover, 2008; Rockliff, Karl, McEwan, Gilbert, Matos, et al., 2011; Gilbert, et al., 2012), anger (Gilbert, Cheung, Irons, & McEwan, 2005; Gilbert & Miles, 2000), paranoid beliefs (Mills, Gilbert, Bellew, McEwan & Gale, 2007) and perfectionism (Gilbert, Durrant, & McEwan, 2006). In contrast, greater self-reassurance is related to better psychological health (Gilbert, et al., 2004; Gilbert, Durrant, & et al., 2006), secure attachment (Irons, et al., 2006) and early memories of warmth and safeness (Richter, et al., 2009).

This scale has also been used in different samples, including: major/severe long-term and complex difficulties in day centre patients (Gilbert & Procter, 2006); in-patients and day-patients from mixed clinical populations (Gilbert, et al., 2010; Judge, Cleghorn, McEwan, & Gilbert, 2012); patients diagnosed with schizophrenia who experienced hostile auditory hallucinations (Mayhew & Gilbert, 2008); depressed patients (Gilbert, McEwan, Catarino & Baião, 2014a).

Given the many studies using the FSCRS original 22-item version as a measure of self-criticism/self-reassurance, more work is required on its psychometric properties and on normative data. This would enable clinicians and researchers to have a better understanding of patients/participants results on the scale. Particularly, by providing for the first time normative data on the scale, we hope to help practitioners to better interpret the level and clinical relevance of the patients' self-criticism.

Therefore, there are three main aims for the present study. The first one is to examine the validity and reliability of the original 22-item scale, in two large samples of the general and clinical population. Based on previous findings, we anticipate good validity and reliability. We also expect that the confirmatory factor analysis supports the original 3-factor model for both populations, differentiating between reassured-self, inadequate-self and hated-self. The second aim of this study is to, for the first time, provide normative data for the interpretation of the results of clinical and nonclinical populations. The third aim is to explore the levels of self-criticism/self reassurance on the two samples, considering the differences by gender and by population. Based on previous studies, we expect that the clinical group reveals higher scores of inadequate-self and hated-self, and lower scores of reassured-self.

Method

Participants

Authors from twelve previous studies using the FSCRS were contacted by email for permission to use the data. Original studies examined subjects including: anhedonia, social rank, defeat and entrapment (Gilbert, Allan, Brough, Melley, et al., 2002), self-criticism and self-warmth/reassurance (Gilbert, Baldwin et al., 2006; Gilbert et al., 2004), rumination (Gilbert et al., 2005), perfectionism, self-criticism, sensitivity to put down, shame (Gilbert et al., 2010; Gilbert, Durrant, et al. 2006.), fears of compassion and happiness (Gilbert et al., 2012, 2014b), self-harm (Gilbert et al., 2010) and Compassionate Mind Training (Gilbert & Procter, 2006; Lucre & Corten 2013; Procter & Bradley, 2005). Seven of the original studies included nonclinical participants (Gilbert, Baldwin, et al., 2006; Gilbert, Durrant et al., 2006; Gilbert et al., 2002, 2004, 2005, 2012; Gilbert & Miles 2000) and five included clinical participants (Gilbert et al., 2010; Gilbert, McEwan, Catarino, & Baião, 2014b; Gilbert & Procter, 2006; Lucre & Corten 2013; Procter & Bradley, 2005). All of the participants completed the questionnaires by hand using pen and paper. Two samples derived from the gathered data: a nonclinical population and a clinical population.

The nonclinical population consists of a total of 887 undergraduate students from a university in the UK (210 males, 676 females). Participants are aged 18-57 years ($Mean = 24.13$; $SD = 7.79$).

The mixed diagnosis clinical sample consists of 171 patients; of those 67 (39.18%) are outpatients, 79 (46.20%) are inpatients, 17 (9.94%) belong to self-help groups and 8 (4.68%) from unknown origin. Regarding diagnosis, depression accounts for the majority of the cases (100 participants, 58.48% of the sample), followed by personality disorder (16 participants, 9.36%), substance abuse (13 participants, 7.60%), anxiety (9 participants, 5.26%), and bipolar disorder (3 participants, 1.54%). For the rest of the participants (30 participants, 17.54% of the sample) information on diagnosis could not be retrieved. Age information is missing in 23 participants (13.5% of the sample), and gender information is missing in 13 participants (7.6% of the sample). Participants with this information are aged 20-69 years ($Mean = 44.22$; $SD = 12.05$), 67 of them are males, and 91 are females.

Measure

Forms of Self-Criticising/Attacking & Self-Reassuring Scale (FSCRS, Gilbert, Clarke-Hempel, Miles and Irons, 2004)

This 22-item scale was developed by Gilbert, Clarke, Hempel, Miles and Irons (2004). Participants answer on a five-point Likert scale ranging from 0 (“not at all like me”), to 4 (“extremely like me”). It measures self-criticism and the ability to self-reassure, when things go wrong for people. There are two forms of self-criticism: *inadequate-self*, which focuses on a sense of personal inadequacy (“I am easily disappointed with myself”); *hated-self*, which measures the desire to hurt or persecute the self (“I call myself names”); These is one factor for being able to *self-reassure* (e.g., “I find it easy to forgive myself”). *Cronbach’s alphas* in nonclinical samples ranged from .89 to .91 for inadequate-self, .82 to .89 for hated-self and .82 to .88 for reassured-self. In clinical samples, *Cronbach’s alphas* ranged from .87 to .89 for inadequate-self, .83 to .86 for hated-self and .85 to .87 for reassured-self.

Analytical Plan

Data analysis was conducted using SPSS (v.20, SPSS Inc. Chicago, IL) for all the descriptive and correlational procedures, and AMOS software (v.20, SPSS Inc. Chicago, IL) for the confirmatory factor analysis.

The existence of outliers was assessed through Mahalanobis distance (D^2) and the normality of variables through coefficients of skewness (Sk) and Kurtosis (Ku).

In order to test the three-factor model of FSCRS (Gilbert et al., 2004; Kupeli et al., 2012), we conducted a confirmatory factor analysis (CFA) in both nonclinical and clinical populations. To test the model, we used Maximum Likelihood estimator, since it is one of the most common estimation methods within this type of statistical procedure (Brown, 2006).

The overall adjustment of the model was assessed by taking into consideration several goodness-of-fit indices, more specifically Chi-Square (χ^2), the Normed Chi-Square (χ^2/df), the Tucker Lewis Index (*TLI*), the Comparative Fit Index (*CFI*) and the Root-Mean Square Error of Approximation (*RMSEA*). In addition to the value of *RMSEA*, *PCLOSE* tests the null hypothesis of *RMSEA* to be no greater than .05. If *PCLOSE* is greater than .05, this means that *RMSEA* is greater than .05 (i.e., the model fit does not have close-fit). The adjustment of the model took into consideration the Modification Indexes (*MI*). In order to test if two different models were significantly

different, we used the chi-square difference test. We've also analyzed the items' factor loadings ($\lambda \geq .50$), since it gives us information regarding the amount of variance in the observed variables that is explained by the underlying construct. In addition, discriminant validity was also examined in order to assess if the latent variable accounted for more variance in the observed variables associated with it than the measurement error (or similar unmeasured influences) or other constructs in the conceptual framework (Fornell & Larcker, 1981). Discriminant validity is obtained by comparing the Average Variance Extracted (AVE) of each factor with the shared variance between factors. The AVE of one factor (A) and the AVE of another factor (B) both need to be larger than their shared variance (i.e., square of the correlation between them) (Hair, Anderson, Tatham & Black, 1998).

As an additional contribution of this study, we also conducted a multi-group CFA, using AMOS software (v.20, SPSS Inc. Chicago, IL), in order to explore and assess if the factor structure of the scale was indeed invariant between both groups. The invariance of the measured model was assessed in both groups by comparing the unconstrained model, measurement weights, structural covariances and measurement residuals. Statistical difference between models were assessed through the difference between Comparative Fit Indices (CFI) (Cheung & Rensvold, 2002).

Reliability of the scale was assessed through Cronbach's alpha (α) and composite reliability (CR), since the latter being less biased and more appropriate for multidimensional scales (Marôco, 2010).

Results

1. Preliminary analysis

1.1 Nonclinical population

Results didn't indicate severe violations of normal distribution ($|Sk| < 3$ and $|ku| < 10$). There were several multivariate outliers, which we have decided not to eliminate from our sample. Dealing with outliers is a rather controversial topic in statistics. Although it has been proposed the elimination of outliers (Marôco & Bispo, 2003), or the transformation of variables (Tabachnick & Fidell, 1989), some authors suggest that they should be kept, since they represent possible observation within general population, thus its results are more generalizable (Hair et al., 1998). Given that this is a sample composed with participants from general population, in which extreme observation are expected to occur, we have decided to keep outliers in this sample.

Nevertheless, we conducted a CFA both with and without the outliers. Results of both analysis showed a better fit model when outliers are not eliminated from our sample.

1.2 Clinical population

Results didn't indicate severe violations of normal distribution ($|Sk| < 3$ and $|ku| < 10$). In this sample, we have found four outliers (observation 169, 23, 2 and 20). Since this sample is composed by participants with a specific psychiatric diagnosis, and since this sample is considerably smaller than the non-clinical sample, keeping extreme observations should have a detrimental impact on data distribution and consequently on results obtained. In addition, we have also conducted CFA analysis with and without outliers, and model fit presented to be better without the outliers. Thus, outliers were eliminated from our clinical sample before proceeding with the analysis. All of the subsequent analysis were performed excluding the outliers (N = 167 patients).

2. Internal consistency

The internal consistency of the FSCRS was calculated for the three subscales (inadequate-self, hated-self and reassured-self) in each of the populations (nonclinical and clinical populations). For the nonclinical population, the *Cronbach's alpha* was .90 for inadequate-self and .85 for both the hated-self and the reassured-self. For the clinical population, *Cronbach's alpha* was .91 for the inadequate-self, .87 for the hated-self and .85 for the reassured-self.

3. Confirmatory Factor Analysis

3.1. Nonclinical population

Model fit indices showed global reasonable fit. Although Chi-square was statistically significant, this indice has been suggested to be greatly influenced by sample size (Jöreskog & Sörbom, 1993), leading to an erroneous conclusion that the model is not fit, when the model is in fact appropriate (Bollen, 1989). A more suitable measure, and a way to minimize the influence of sample size, is by using Normed Chi-Square, which should be between 2 and 5 (Bollen, 1989; Marsh and Hocevar, 1985; Tabachnick, & Fidell, 2007). CFI and TLI reach the suggested cut-off value of .90 (Marôco, 2010). RMSEA has been regarded as one of the most informative fit indices (Diamantopoulos & Sigaw, 2000). Although RMSEA was between .05 and .08 (Hu & Bentler, 1998), some concerns were raised by looking into its PCLOSE (see Table 1).

----- insert Table 1 here -----

The modification indices showed that item 22 (“I do not like being me”), which loads onto HS factor, could also be predicted by RS factor. In fact, similar results concerning item 22 have also occurred elsewhere (Kupeli et al., 2012). In order to deal with these results, we tested two models: one in which item 22 is predicted both by HS and RS (Model 2); and a simplified one, with item 22 deleted from the model (Model 3). Although both models were improved (Model 2: DIFFTEST; $\Delta\chi^2 = 72.6$, $df = 1$; Model 3: DIFFTEST; $\Delta\chi^2 = 179.3$, $df = 20$), PCLOSE was still $\leq .001$, which indicates a poor close fit.

Our modification indexes also suggested that some items’ errors should correlate (items 1 and 2, 3 and 5, 8 and 22, 9 and 10, and 15 and 18), and for that reason we tested a model in which we correlated the errors associated with those items (Model 4), and maintained item 22 saturating only in HS, as proposed by the original authors (Gilbert et al., 2004). In fact, this model showed the best fit (see Table 1) and was significantly better than the original model (DIFFTEST; $\Delta\chi^2 = 351.304$, $df = 5$) (see Figure 1).

----- insert Figure 1 here -----

Since it has been previously suggested the possibility of combining HS and IS in a global factor of *self-criticism*, we decided to also test a two-factor model. Fit indices results showed poor fit of the model (Model 5).

Our results suggested good composite reliability (.94 for Inadequate-Self, .90 for Hated-Self and .87 for Reassured-Self), with coefficients of determination (R^2) ranging between .25 and .64. The calculated AVEs was .62 for Inadequate-Self, .65 for Hated-Self and .48 for Reassured-Self. Discriminant validity was assessed by comparing AVE and square correlation between factors. The calculation of squared correlations between Reassured-Self and Inadequate-Self ($r^2 = .36$), Inadequate-Self and Hated-Self ($r^2 = .60$) and Reassured-Self and Hated-Self ($r^2 = .46$), when compared to respective AVEs, suggest a good discriminant validity between all three factors.

3.2. Clinical population

Results show that the three-factor model has a reasonable fit [$\chi^2 = 332.292$, $p \leq .001$; $\chi^2/d.f. = 1.613$; $CFI = .936$, $TLI = .929$, $RMSEA = .061$ (CI = .048, .073, $p = .073$)] (see Table 2).

----- insert Table 2 here -----

However, by considering the modification indexes values, it is suggested a covariance between errors of items 8 (“I still like being me”) and 11 (“I can still feel lovable and acceptable”). In addition, it is also suggested that item 14 (“*I find it difficult to control my anger and frustration at myself*”) loads both in factors IS and HS. For that reason, we conducted a CFA with a model in which we correlate the errors, and in which both HS and IS predict item 14 (Model 2); a model in which we correlate the errors and eliminate item 14 from the model (Model 3); and a model in which we only correlated the errors (Model 4). The latter presented significantly better goodness-of-fit indices when compared with the initial model (DIFFTEST; $\Delta\chi^2 = 15.156$, $df = 1$) (see Figure 2).

----- insert Figure 2 here -----

Regarding the construct validity, our results also suggest FSCRS presents a very good composite reliability in our clinical sample, with .95 for Inadequate-Self, .91 for Hated-Self and .91 for Reassured-Self, with coefficients of determinations between .37 and .74. Our results showed AVEs of .67 for Inadequate-Self, .66 for Hated-Self and .58 for Reassured-Self ($> .50$). Good discriminant validity was obtained between Reassured-Self and Inadequate-Self ($r^2 = .42$), and between Reassured-Self and Hated-Self ($r^2 = .42$). However, discriminant validity was less evident between Inadequate-Self and Hated-Self ($r^2 = .79$), as has previously occurred (Castilho et al., 2013).

Since our results suggested a high correlation between factor Inadequate-Self and Hated-Self, and given that a two-factor model of FSCRS (with Inadequate-Self and Hated-Self as a global *Self-Criticism* factor) has been previously tested (Kupeli et al., 2012), we decided to also test a two-factor model of FSCRS (Model 5) with our clinical sample. Results showed poor adjustment (see Table 2).

3.3. Multi-group CFA for clinical and nonclinical populations

In addition to the estimation of the fit of different models separately both for clinical and nonclinical populations, we conducted a multi-group CFA in order to test the measurement invariance of the model in both populations.

----- insert Table 3 here -----

The baseline unconstrained model tested the factor structure of FSCRS simultaneously across clinical and nonclinical populations, with no constraints imposed. This model presented good model fit indices of $\chi^2 = 1384,454$, $p \leq .001$; $\chi^2/\text{d.f.} = 3,360$; CFI= .914; TLI= .903; RMSEA= .047, $p = .934$.

The measurement weights tested the invariance of the factor loadings across the two samples by containing equality on these parameters. The model showed fit indices of $\chi^2 = 1431,097$, $p \leq .001$; $\chi^2/\text{d.f.} = 3,320$; CFI= .911; TLI= .905; RMSEA= .047, $p = .963$. Given that the χ^2 difference test is highly sensitive to sample size, Cheung and Rensvold (2002) suggested using ΔCFI as an alternative for measuring invariance between groups. A ΔCFI value higher than .01 is indicative of a significant drop in fit, i.e., a non-equivalence between groups. The $\Delta\text{CFI} = -.003$ suggested that equality constraints of factor loadings did hold across the two populations.

The structural covariances tested (i.e., a model in which both the factor loadings and covariances are fixed) showed model fit indices of $\chi^2 = 2026,617$, $p \leq .001$; $\chi^2/\text{d.f.} = 4,415$; CFI= .861; TLI= .860; RMSEA= .057, $p \leq .001$. The ΔCFI between the structural covariances and measurement weights was -.05, which also confirmed the invariance between groups.

Finally, a measurement residuals model (i.e. with factor loadings, covariances and residuals fixed) was tested, and showed fit indices of $\chi^2 = 2243,484$, $p \leq .001$; $\chi^2/\text{d.f.} = 4,664$; CFI= .844; TLI= .850; RMSEA= .059, $p \leq .001$. The difference in CFI between measurement residuals and structural covariances also suggested the invariance of the structural model across the two populations.

This data suggest that the original structure of FSCSR is in fact fit to assess forms of self-criticism and self-reassurance, both in clinical samples and in samples composed by participants from general nonclinical population.

4. Normative study and group comparisons

4.1. Nonclinical and clinical population

4.1.1 Normative Study

Based on the original factor structure confirmed above, normative data is presented. For the nonclinical (N = 887) and clinical (N = 167) populations, descriptive data was collected and is displayed on Table 4.

----- insert Table 4 here -----

4.1.2 Comparisons between clinical and nonclinical populations

Means and standard deviations of the clinical and nonclinical populations were compared to test for significant differences using an independent measures *t* test. Results showed a significant difference between the two populations in relation to reassured-self ($t(1048) = -19.32, p = .000$), inadequate-self ($t(248.703) = 15.13, p = .000$) and hated-self scores ($t(209.216) = 18.02, p = .000$). For reassured-self, patients reported lower scores than non patients ($M = 10.68; SD = 6.51$ compared to $M = 20.27; SD = 5.77$). For inadequate-self, patients reported higher scores than non patients ($M = 27.47; SD = 7.51$ compared to $M = 17.72; SD = 8.29$), and the same for the hated-self scores ($M = 12.26; SD = 5.67$ compared to $M = 3.88; SD = 4.59$).

4.2. Normative data by gender

4.2.1. Nonclinical population

For the nonclinical population, descriptive data on 210 male participants and 676 female participants is displayed in Table 5.

----- insert Table 5 here -----

4.2.2. Clinical population

For the clinical population, descriptive data on 64 male and 91 female mixed diagnosis participants was also collected, as shown in Table 6.

----- insert Table 6 here -----

4.2.3. Comparisons between genders in clinical and nonclinical populations

No significant gender differences were found in the clinical population.

In the nonclinical population, males and females had significant differences on the three subscales. Men ($M = 21.20$; $SD = 5.27$) scored significantly higher than women ($M = 19.98$; $SD = 5.90$) in reassured-self ($t(386,936) = 2.82$; $p = .005$). On the other hand, men scored lower ($M = 16.42$; $SD = 7.44$) than women ($M = 18.11$; $SD = 8.50$) in inadequate-self ($t(880) = -2.59$; $p = .010$); and also marginally lower ($M = 3.36$; $SD = 3.71$) than women ($M = 4.05$; $SD = 4.83$) in hated-self ($t(879) = -1.90$; $p = .058$).

Discussion

This study shows that the original FSCRS is a robust and reliable measure of self-criticism and its contrast, self reassurance. The three subscales of the FSCRS show good reliability, either in the nonclinical population and the clinical population.

The **confirmatory factor analysis** supports the three-factor solution obtained by the authors during original 22-item scale development and validation (Gilbert et al., 2004) both in the nonclinical and clinical populations, confirming that the three-factor model of FSCRS is a well-adjusted measure for assessing the two forms of self-criticism and a form of self-reassurance. As hypothesized, even in clinically diverse populations expected to present different levels of self-criticism, the three forms stand as independent. This suggests that self-criticism is a process, which is not only transdiagnostical, but also present in people with and without psychopathology (in different levels). The invariance of the multi-group analysis also confirms the original structure.

This result is in line with previous studies (eg. Castilho et al., 2013 ; Kupeli, et al., 2013). It highlights the fact that self-criticism shouldn't be seen as one single dimension but as having different forms and functions, that may operate differently, have different originators, and respond to different types of therapy.

We also present **normative data** for each population (clinical and nonclinical) and for each gender within each population based on Means, Standard Deviations, Medians and Percentiles. It is unknown if this normative data would be represented in different cultural populations or age groups, which is a limitation of this study. It's also possible that since the majority of the clinical groups were depressed, other forms of psychopathology may have slightly different loadings. Being the populations from this study a collection of previous samples, some of the demographic information could not be standardized for all of the participants, which prevented deeper study of the data.

Nonetheless these data can aid clinicians and researchers' interpretation of self-criticism results, which is essential given its impact on mental health (Gilbert & Irons, 2005; Zuroff, Santor, & Mongrain, 2005) and on treatment response (Cox et al., 2002; Rector, et al. 2000).

Examining the **differences between populations**, as expected there are significant differences on the three subscales of the FSCRS, with the nonclinical population scoring higher on reassured-self and lower on inadequate-self and hated-self. This finding is in line with previous findings which report a link between self-criticism and psychopathological traits and diagnosis (Gilbert, et al., 2004; 2005; Gilbert, Durrant, et al., 2006; Gilbert & Irons, 2005; Gilbert & Miles, 2000; Mills, et al., 2007; Zuroff et al., 2005). Again, this link between self-criticism and a wide range of psychopathology suggests that self-criticism should be considered as a process and transdiagnostical trait, better than a simple symptom.

As mentioned, the self-hating dimension might represent a more pathological domain of self-criticism, associated with self-harm and more borderline phenomenology (Gilbert, et al., 2004; Gilbert, et al., 2010). In this study, since the great majority of the clinical sample was diagnosed with depression, it was not possible to test the link between severity of the psychopathology and inadequate vs. hated-self scores. This phenomenological difference between the inadequate and the hated forms of self-criticism remains in need of more in depth research.

In terms of **gender differences**, in the nonclinical population men scored significantly higher on reassured-self and lower on inadequate-self and hated-self than women, suggesting that generally women are more self-critical and less self-reassuring. However, in the clinical population there were no significant gender differences. It can be that as individuals become depressed or mentally unwell the same processes are operating in both men and women. This is in line with an earlier study with depressed participants which found no gender differences in the correlation between depression and internalization (related to self-criticism) (Gilbert, Irons, Olsen, Gilbert, & McEwan, 2006).

Insights into the forms and functions as well as the origins and treatment of self-criticism will continue to rise with the development of new measures. Whether other dimensions emerge, beyond those of feeling inadequate and wanting to self-correct or self-hating, awaits future work. What this study does confirm however is that the three dimensions identified in the FSCRS stand both in clinical and nonclinical samples, even

using a robust analysis as the CFA (Curran, West & Finch, 1996). This represents an important addition to the scale validation as well as to its use in clinical and research settings.

REFERENCES

- Bollen, K. A. (1989). *Structural Equations with latent variables*. New York: Wiley.
- Brown, T. A. (2006). *Confirmatory factor analysis for applied research*. New York: Guilford Press.
- Castilho, P., Pinto-Gouveia, J. & Duarte, J. (2013). Exploring self-criticism: Confirmatory Factor Analysis of the FSCRS in a clinical and in a nonclinical sample. *Clinical Psychology & Psychotherapy*. DOI: 10.1002/cpp.1881.
- Cheung, G. W., & Rensvold, R. B. (2002). Evaluating goodness-of-fit indexes for testing measurement invariance. *Structural Equation Modeling*, 9, 233–255. DOI:10.1207/S15328007SEM0902_5.
- Cox, B. J., Walker, J.R., Enns, M.W., Karpinski, D.C. (2002). Self-criticism in generalized social phobia and response to cognitive-behavioral treatment. *Behaviour Therapy*, 33, 479–491. DOI: 10.1016/S0005-7894(02)80012-0.
- Curran, P. J., West, S. G., & Finch, G. F. (1996). The robustness of test statistics to nonnormality and specification error in confirmatory factor analysis. *Psychological Methods*, 1, 16–29. DOI: 10.1037/1082-989X.1.1.16.
- Diamantopoulos, A. & Siguaw, J.A. (2000), *Introducing LISREL*. London: Sage Publications.
- Fornell, C. & Larcker, D.F., (1981). Evaluating structural equation models with unobservable variables and measurement error. *Journal of Marketing Research*, 18 (1), 39-50.
- Gilbert, P. (2010). *Compassion Focused Therapy: The CBT Distinctive Features Series*. London: Routledge.
- Gilbert, P., Allan, S., Brough, S., Melley, S. & Miles, J.N.V. (2002). Relationship of anhedonia and anxiety to social rank, defeat and entrapment. *Journal of Affective Disorders*, 71, 141–151. DOI: 10.1016/S0165-0327(01)00392-5.

- Gilbert, P., Baldwin, M.W., Irons, C., Baccus, J.R. & Palmer, M. (2006). Self-criticism and self-warmth: An imagery study exploring their relation to depression. *Journal of Cognitive Psychotherapy*, 20, 183-200. DOI: 10.1891/jcop.20.2.183.
- Gilbert, P., Cheung, M., Irons, C., & McEwan, K. (2005). An exploration into depression-focused and anger-focused rumination in relation to depression in a student population. *Behavioural and Cognitive Psychotherapy*, 33, 273-283. DOI: 10.1017/S1352465804002048.
- Gilbert, P., Clarke, M., Hempel, S., Miles, J.N.V., & Irons, C. (2004). Criticizing and reassuring oneself: An exploration of forms, styles and reasons in female students. *British Journal of Clinical Psychology*, 43, 31-50. DOI: 10.1348/014466504772812959.
- Gilbert, P., Durrant, R., & McEwan, K. (2006). Investigating relationships between perfectionism, forms and functions of self-criticism, and sensitivity to put-down. *Personality and Individual Differences*, 41, 1299-1308. DOI: 10.1016/j.paid.2006.05.004.
- Gilbert, P., & Irons, C. (2005). Focused therapies and compassionate mind training for shame and self-attacking. In P. Gilbert (Ed.), *Compassion: Conceptualisations, research and use in psychotherapy* (pp. 263–325). London: Routledge.
- Gilbert, P., Irons, C., Olsen, K., Gilbert, J & McEwan, K (2006) Interpersonal sensitivities: Their link to mood, anger and gender. *Psychology and Psychotherapy: Theory Research and Practice*, 79, 37 -51. DOI: 10.1348/147608305X43856.
- Gilbert, P., McEwan, K., Catarino, F., & Baião, R. (2014a). Fears of Negative Emotions in Relation to Fears of Happiness, Compassion, Alexithymia and Psychopathology in a Depressed Population: A Preliminary Study. *Journal of Depression & Anxiety*, DOI: 10.4172/2167-1044.S2-004.
- Gilbert, P., McEwan, K., Catarino, F., & Baião, R. (2014b). Fears of compassion in a depressed population - Implications for Psychotherapy. *Journal of Depression & Anxiety*. DOI: 10.4172/2167-1044.S2-00.
- Gilbert, P., McEwan, K., Gibbons, L., Chotai, S., Duarte, J. & Matos, M. (2012). Fears of compassion and happiness in relation to alexithymia, mindfulness and self-criticism. *Psychology and Psychotherapy: Theory, Research and Practice*, 8, 374-390. DOI: 10.1111/j.2044-8341.2011.02046.x.

- Gilbert, P., McEwan, K., Irons, C., Bhundia, R., Christie, R., Broomhead, C., & Rockliff, H. (2010). Self-harm in a mixed clinical population: The roles of self-criticism, shame, and social rank. *British Journal of Clinical Psychology, 49*, 563- 576. DOI: 10.1348/014466509X479771.
- Gilbert, P. & Miles, J. N. V. (2000). Sensitivity to social put-down: Its relationship to perceptions of social rank, shame, social anxiety, depression, anger and self-other blame. *Personality and Individual Differences, 29*, 757–774. DOI: 10.1016/S0191-8869(99)00230-5.
- Gilbert, P. & Procter, S. (2006). Compassionate mind training for people with high shame and self-criticism: A pilot study of a group therapy approach. *Clinical Psychology and Psychotherapy, 13*, 353-379. DOI: 10.1002/cpp.507.
- Goodall, T., Gilbert, P. & McEwan, K. (Submitted). Early Memories of Childhood Warmth: The Effect on Self-Evaluation and Future Emotions.
- Hair, J.F., Anderson, R.E., Tatham, R.L. & Black, W.C. (1998). *Multivariate data analysis*. New York: Prentice Hall.
- Hu, L. & Bentler, P. (1998). Fit indices in covariance structure modeling: Sensitivity to underparameterized model misspecification. *Psychological Methods, 3*, 424-453. DOI: 10.1037/1082-989X.3.4.424.
- Irons, C., Gilbert, P., Baldwin, M.W., Baccus, J.R., & Palmer, M. (2006). Parental recall, attachment relating and self-attacking/self-reassurance: Their relationship with depression. *British Journal of Clinical Psychology, 45*, 297-308. DOI: 10.1348/014466505X68230.
- Jöreskog, K. & Sörbom, D. (1993). *LISREL 8: Structural Equation Modeling with the SIMPLIS Command Language*. Chicago, IL: Scientific Software International Inc.
- Judge, L., Cleghorn, A., McEwan, K & Gilbert, P. (2012). An exploration of group-based compassion focused therapy for a heterogeneous range of clients presenting to a community mental health team. *International Journal of Cognitive Therapy, 5*, 420–429. DOI: 10.1521/ijct.2012.5.4.420.
- Kupeli, N., Chilcot, J., Schmidt, U. H., Campbell, I. C. and Troop, N. A. (2013), A confirmatory factor analysis and validation of the forms of self-criticism/reassurance scale. *British Journal of Clinical Psychology, 52*, 12–25. DOI: 10.1111/j.2044-8260.2012.02042.x.

- Longe, O., Maratos, F.A., Gilbert, P., Evans, G., Volker, F., Rockliffe, H. & Rippon G. (2010). Having a word with yourself: Neural correlates of self-criticism and selfreassurance. *Neuroimage*, 49, 1849-1856. DOI: 10.1016/j.neuroimage.2009.09.019.
- Lucre, K.M. & Corten, N. (2013). An exploration of group compassion-focused therapy for personality disorder. *Psychol Psychother*, 86(4), 387-400. DOI: 10.1111/j.2044-8341.2012.02068.x.
- Marôco, J. (2010). *Análise estatística com o PASW Statistics*. Pêro Pinheiro: ReportNumber.
- Marôco, J. & Bispo, R. (2003). *Estatística aplicada às ciências sociais e humanas*. Lisboa: Climepsi Editores.
- Marsh, H. W. & D. Hocevar (1985). Application of confirmatory factor analysis to the study of self-concept: First- and higher-order factor models and their invariance across groups. *Psychological Bulletin*, 97, 562–582. DOI: 10.1037/0033-2909.97.3.562.
- Mayhew, S.L. & Gilbert, P. (2008). Compassionate Mind Training with People Who Hear Malevolent Voices: A Case Series Report. *Clinical Psychology and Psychotherapy*, 15, 113-138. DOI: 10.1002/cpp.566.
- Mills, A., Gilbert, P., Bellew, R., McEwan, K. & Gale, C. (2007). Paranoid Beliefs and Self-Criticism in Students. *Clinical Psychology and Psychotherapy*, 14, 358–364. DOI: 10.1002/cpp.537.
- Neff, K. D. (2003). Self-compassion: an alternative conceptualization of a healthy attitude toward oneself. *Self & Identity*, 2, 85–102. DOI:10.1080/15298860309032.
- Procter, S. & Bradley, B. (2005). [Compassionate Mind Training in a clinical sample]. Unpublished raw data.
- Rector, N.A., Bagby, R.M., Segal, Z.V., Joffe, R.T. & Levitt, A. (2000). Self-criticism and dependency in depressed patients treated with cognitive therapy or pharmacotherapy. *Cognitive Therapy and Research*, 24, 571–584. DOI: 10.1023/A:1005566112869.
- Richter, A., Gilbert, P., & McEwan, K. (2009). Development of an early memories of warmth and safeness scale and its relationship to psychopathology. *Psychology and Psychotherapy: Theory, Research and Practice*, 82, 171-184. DOI: 10.1348/147608308X395213.

- Rockliff, H., Gilbert, P., McEwan, K., Lightman, S., & Glover, D. (2008). A pilot exploration of heart rate variability and salivary cortisol responses to compassion focused imagery. *Journal of Clinical Neuropsychiatry*, *5*, 132–139.
- Rockliff, H., Karl, A., McEwan, K., Gilbert, J., Matos, M., & Gilbert, P. (2011). Effects of Intranasal Oxytocin on 'Compassion Focused Imagery'. *Emotion*, *11*, 1388-96. DOI: 10.1037/a0023861.
- Tabachnick, B.G. & Fidell, L.S. (1989). *Using multivariate statistics*. New York: Harper Collins.
- Whelton, W.J. & Greenberg, L.S. (2005). Emotion in self-criticism. *Personality and Individual Differences*, *38*, 1583–1595. DOI: 10.1016/j.paid.2004.09.024.
- Zuroff, D.C., Santor, D., & Mongrain, M. (2005). Dependency, self-criticism, and maladjustment. In J.S. Auerbach, K.N. Levy, & C.E. Schaffer (Eds.). *Relatedness, self-definition and mental representation. essays in honour of Sidney J. Blatt*. (pp. 75–90). London: New York.

TABLES

Table 1. Factor structure of the FSCRS in a nonclinical population (N = 887)

		χ^2	p	$\chi^2/d.f.$	CFI	TLI	RMSEA	PCLOSE
1	3-factor FSCRS	1052.710	$\leq .001$	5.110	.909	.898	.068	$\leq .001$
2	22 in RS and HS	980.085	$\leq .001$	4.781	.917	.906	.066	$\leq .001$
3	22 removed	873,384	$\leq .001$	4.696	.920	.909	.065	$\leq .001$
4	Correlated errors	701,406	$\leq .001$	3,490	.946	.938	.053	.104
5	2-factor FSCRS	1632,596	$\leq .001$	7,849	.847	.830	.088	$\leq .001$

Figure 1. . Item loading of the FSCRS in a nonclinical population (N = 887)

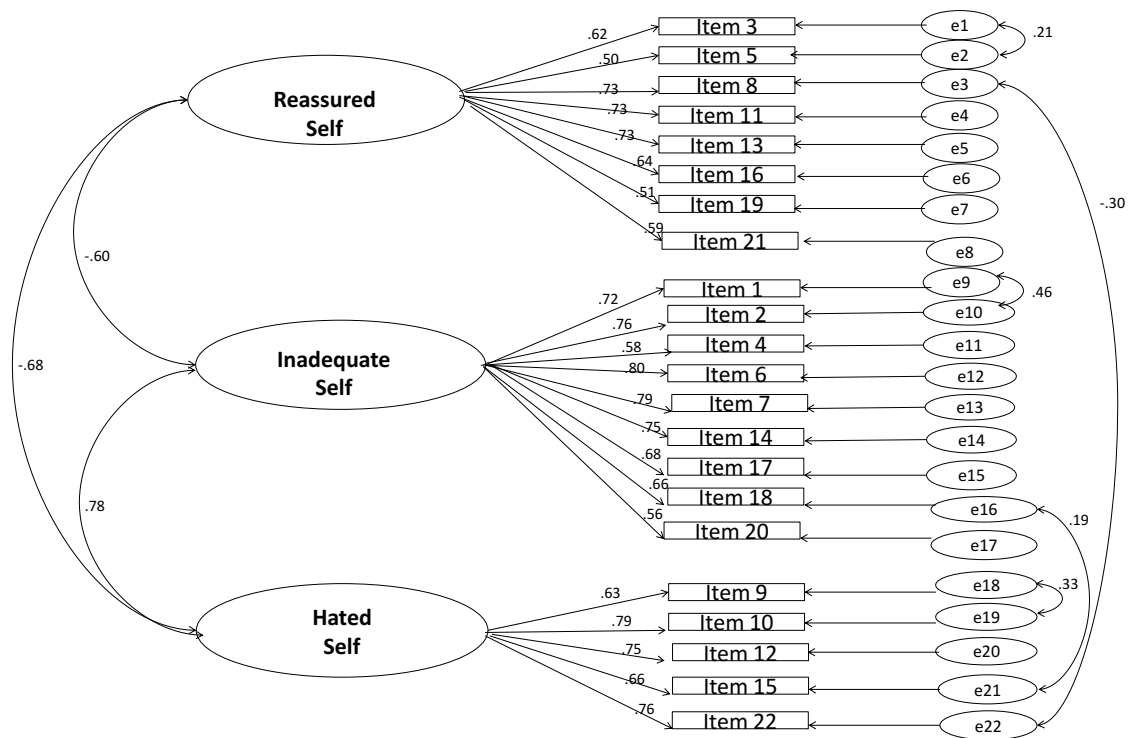


Table 2. Factor structure of the FSCRS in a clinical population (N = 167)

		χ^2	<i>p</i>	$\chi^2/d.f.$	<i>CFI</i>	<i>TLI</i>	<i>RMSEA</i>	<i>PCLOSE</i>
1	3-factor FSCRS	332.292	≤ .001	1.613	.936	.929	.061	.073
2	14 in IS and HS	316.587	≤ .001	1.552	.943	.936	.058	.153
3	14 removed	306.668	≤ .001	1.657	.936	.927	.063	.048
4	Correlated errors	317.136	≤ .001	1.547	.940	.936	.057	.161
5	2-factor FSCRS	365.037	≤ .001	1.755	.921	.912	.067	.008

Figure 2. Item loading of the FSCRS in a clinical population (N = 167)

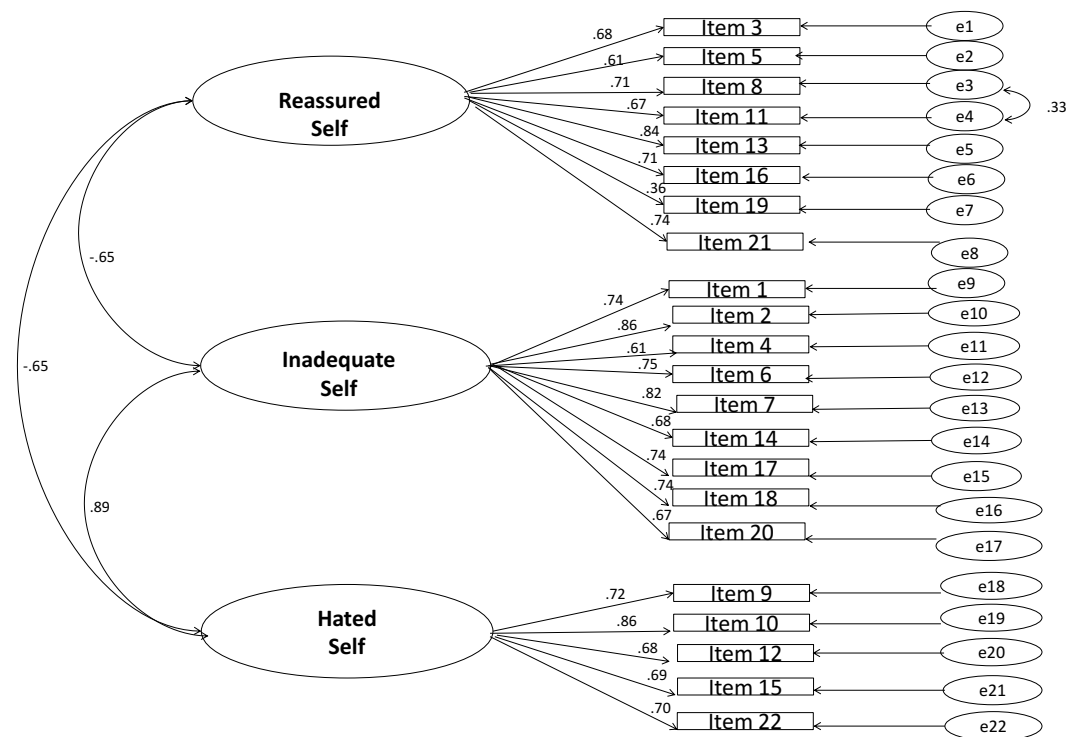


Table 3.

		χ^2	<i>p</i>	$\chi^2/d.f.$	<i>CFI</i>	<i>TLI</i>	<i>RMSEA</i>	<i>PCLOSE</i>
1	Unconstrained	1384,454	≤ .001	3,360	.914	.903	.047	.934
2	Measurements	1431,097	≤ .001	3,320	.911	.905	.047	.963
3	Structural covariances	2026,617	≤ .001	4,415	.861	.860	.057	≤ .001
4	Measurements residuals	2243,484	≤ .001	4,664	.844	.850	.059	≤ .001

Table 4. Means, Standard Deviations, Median and Percentiles of the FSCRS in nonclinical (N = 887) and clinical (N = 167) populations

		Inadequate-self		Hated-self		Reassured-self	
		Nonclinical Male	Nonclinical Female	Nonclinical Male	Nonclinical Female	Nonclinical Male	Nonclinical Female
Mean		16.42	18.11	3.36	4.05	21.20	19.98
(SD)		(7.44)	(8.50)	(3.71)	(4.83)	(5.27)	(5.90)
Percentiles	5%	13.00	4.00	.00	.00	4.00	10.00
	25%	18.00	11.00	.00	.00	11.00	16.00
	50% (Median)	21.00	18.00	2.00	2.00	16.00	20.00
	75%	25.00	24.00	5.00	6.00	22.00	24.00
	95%	29.00	33.00	12.00	15.00	28.45	30.00

Table 5. Means, Standard Deviations, Median and Percentiles of the FSCRS for male (N = 210) and female (N = 676) nonclinical population

		Inadequate-self		Hated-self		Reassured-self	
		Non-clinical	Clinical	Non-clinical	Clinical	Non-clinical	Clinical
Mean		17.72	27.47	3.88	12.26	20.27	10.66
(SD)		(8.29)	(7.51)	(4.59)	(5.67)	(5.77)	(6.51)
Percentiles	5%	4.00	12.60	.00	1.00	10.00	1.00
	25%	11.00	23.00	.00	8.00	16.00	6.00
	50% (Median)	18.00	30.00	2.00	13.00	21.00	10.00
	75%	24.00	34.00	6.00	16.00	24.00	14.00
	95%	32.00	36.00	14.00	20.00	29.95	22.60

Table 6. Means, Standard Deviations, Median and Percentiles of the FSCRS in a male (N = 64) and female (N = 91) clinical population

		Inadequate-self		Hated-self		Reassured-self	
		Clinical Male	Clinical Female	Clinical Male	Clinical Female	Clinical Male	Clinical Female
Mean		26.61	27.51	12.13	11.91	10.66	11.13
(SD)		(7.19)	(7.89)	(5.19)	(6.10)	(4.72)	(7.51)
Percentiles	5%	12.25	11.60	1.00	.60	2.25	1.00
	25%	23.00	22.00	8.25	7.00	7.25	4.00
	50% (Median)	28.50	30.00	13.00	13.00	11.00	10.00
	75%	32.00	34.00	16.00	17.00	13.00	17.00
	95%	36.00	36.00	19.00	20.00	20.00	25.60