Masterthesis

Psychometric Properties of the Forms of Self-Criticizing/-Attacking and Self-Reassuring Scale in People with Eating Disorders in the Netherlands

Nora Becker

S1214705

Positive Psychology and Technology

Universiteit Twente

Supervision: Peter ten Klooster

Sander de Vos

22.01.2018

Abstract

Self-criticism is a key predictor of eating disorder symptoms. Therefore, a valid and reliable instrument is needed in order to assess self-criticism and its counterpart self-reassurance. The Forms of Self-Criticizing / - Attacking and Self-Reassuring Scale (FSCRS; Gilbert, Clarke, Hempel, Miles, & Irons, 2004) was examined for its factorial validity, internal consistency and convergent validity within a sample of Dutch patients with eating disorders (N=502). The results show that the intended three-factor model with the components of *inadequate self, hated self* and *reassured self* fits the data best. Although the factor model had to allow two correlations of item residual variances, the expected factor validity was confirmed. Furthermore, the research found good internal consistency for the three scales and good convergent validity. The expected strong negative correlation of the *inadequate self* and *hated self* with mental wellbeing was confirmed, whereas *reassured self* had a strong positive relationship with mental wellbeing just as expected. Although validity and reliability were supported with the present study, further examinations such as incremental and concurrent validity, measurement invariance testing and reliability testing are needed to further confirm the statement of the FSCRS having good psychometric properties.

Content

1 Introduction

Whenever people are struggling with setbacks and failures, they are using various forms of self-to-self relating. In other words, everybody has a different strategy to treat themselves when things go wrong. People who tend to treat themselves with self-critical thinking, hold constantly negative beliefs about themselves, treat themselves with harsh self- judgement and do not experience any pleasure in achieving their goals (Gilbert, Durrant, & McEwan, 2006). Self-criticism is defined as a maladaptive strategy for emotion regulation: "a persistent tendency for negative self-evaluation that instils feelings of shame and low self-worth" (Falconer, King, & Brewin, 2015; p.352). Specifically, it refers to the extent towards which individuals perceive or evaluate themselves as substandard or inferior in relation to others (Mongrain, Vettese, Shuster, & Kendal 1998).

Research shows that self-criticism presents in various forms (i.e. what people do and say in face of failure) and functions (i.e. what people's reasons are for being self-critical in situations of setbacks; Falconer, King, & Brewin, 2015; Gilbert, Durrant, & McEwan, 2006; Gilbert et al., 2004). According to Gilbert, Durrant and McEwan (2006) self-criticism is comprised out of two components: one form concentrates on feelings of disappointment and inadequacy, the other form focusses on self-disgust and self-hating.

As an adversary strategy to deal with failure, Gilbert et al. (2004) explain the concept of self-reassuring as the 'ability to focus on one's positives and be reassuring to oneself when things go wrong' (p. 35). Reassuring enhances a warm attitude towards oneself and compassion and acceptance for failures. Palmeira, Pinto-Gouveia, Cunha, and Carvalho (2017) describe self-reassurance, a component of self-compassion, as a crucial process to protect against psychological malfunctions. Also, contrary to self-criticism, self-reassurance as well as self-compassion enhance psychological well-being and are associated with good mental health (Sommers-Spijkerman, Trompetter, ten Klooster, Schreurs, Gilbert, &

Bohlmeijer, 2017). Research suggests that self-criticism is strongly associated with the presence of depression and anxiety (Gilbert, Clarke, Hempel, Miles, & Irons, 2004), self-harm (Gilbert et al., 2010), anger (Gilbert, Cheung, Irons, & McEwan, 2005; Gilbert & Miles, 2000), paranoid beliefs (Mills, Gilbert, Bellew, McEwan, & Gale, 2007), and perfectionism (Gilbert, Durrant, & McEwan, 2006). In conclusion, prominent levels of self-critical behavior and a lack of self-reassurance could impose a significant risk for the onset of psychopathology and decreases in mental health or wellbeing.

Self-criticism does not only influence the onset of psychopathology but also affects the effectiveness of treatment outcomes. Research shows that self-criticism is positively related to a hostile mood state during therapy (Kannan & Levitt, 2013) as well as to poor response in cognitive behavioral therapy (Rector, Bagby, Segal, Joffe, & Levitt, 2000). When self-criticism is just seen as a symptom of psychopathology, and therefore not treated specifically, the therapy outcome will be less effective. Rector et al. (2000) stated that whenever self-critical behavior is focused upon in cognitive therapy, the treatment of the associated psychopatholgy will be more successful. In a study, assessing if cognitive therapy is producing a change in self-criticism scores, Rector et al. (2000) found that patients with a higher score on self-criticism did respond poorer to the treatment. However, patients with pretreatment for self-criticism reported improved treatment response. Gilbert and Procter (2006) also describe the importance of specific treatment of self-criticism. So far, Compassionate Mind Training (CMT) is the only therapeutic process, based on evidence-based techniques (Cognitive Behavioral Therapy and Dialectical Behavioral Therapy; Gilbert & Procter, 2006) that is specifically offered for people with high levels of self-criticism and shame. In their study, Gilbert and Procter (2006) included persons with major long term and complex difficulties who were in current treatment in a day center to their study.

It is therefore clear that self-criticism plays an important role in the therapeutic process, especially when the focus is on the regulation of negative affect or mood disorders (Rector et al., 2000). The role of self-criticism and its impact in relation to the different types of emotional disorders have received the benefit of more than five decades of empirical research, yet the impact of self-criticism on other forms of pathology such as eating disorders, less so. Therefore, it is important to take the impact of self-criticism on eating disorders into focus.

Eating disorders are categorized by the Diagnostic and Statistical Manual of Mental Disorders (DSM-5; APA, 2013) in several diagnoses, whereby the present study focusses on the most prevalent ones: anorexia nervosa (AN), bulimia nervosa (BN), binge-eating disorder (BED) and Other Specified Feeding-/Eating Disorders (OSFED). All four are characterized with disturbed eating or eating-related behavior which results in abnormal preoccupation with the consumption of food which results in significant impairment of physical health. Overevaluation of body weight and shape is in both AN and BN seen as essentially for the diagnosis (Duarte, Ferreira, & Pinto-Gouveia, 2016). The way AN and BN patients handle their weight(-loss) however differs: AN-patients selectively restrict the intake of food whereas BN-patients suffer from repeated binges (uncontrolled intake of unusually enormous amounts of food with the sense of aversion) followed by compensatory behavior such as self-induced vomiting or laxative misuse (APA, 2013; Fairburn & Harrison, 2003).

Prevalence of these disorders spreads predominantly in adolescent and young adult females from Western societies, with a lifetime prevalence of 0.9% and 1.5% respectively (Hudson, Hiripi, Pope, & Kessler, 2007). So far, little is known about the prevalence and diagnosis of BED. It is characterized as frequent binge eating without any weight-controlling behavior (APA; 2013; Fairburn & Harrison, 2003) and has no further commonalities with the stated eating disorders. Age of onset and gender preoccupation tend to differ from those of

AN and BN. BED-patients typically present at the age of 40 (Fairburn & Harrison, 2003) and are more equally distributed in gender (Smink, van Hoeken, & Hoek, 2012). The exact prevalence of BED has not been established so far, since most of the BED patients were diagnosed within treatment for obesity. The diagnosis of OSFED is stated when patients do not meet all criteria for one of the other stated eating disorders (APA, 2013). The diagnosis of OSFED has a prevalence of 1 to 1.4% (Hay, Girosi, & Mond, 2015).

The underlying characteristics of people with different types of eating disorders are similar: patients with eating disorders have a high sense of inferiority and harsh self-judgement and they tend to be more critical and attacking towards themselves in phases of failure (Dunkley & Grilo, 2007). It has been found that self-criticism is a profound and independent predictor of symptoms of eating disorders (Fennig et al., 2008), as the symptoms of eating-related behavior appear more often and more severe when faced with setbacks.

Thew, Gregory, Roberts and Rimes (2017) confirmed in their study that self-criticism in eating disorders is experienced at high levels and in the same severity as in depression. Both these clinical groups report significantly more frequent, persistent and uncontrollable self-criticism compared to non-clinical samples. Additionally to that finding, a previous study by Duarte, Ferreira and Pinto-Gouveia (2016) investigated whether patients with different types of eating disorders reported different levels of self-criticism. They found that patients of several types of eating disorders showed similar levels of self-criticism.

The attention for the concept of self-criticism and its role in psychopathology continuously grows, and empirical research on that matter becomes more frequent. Therefore, a reliable and valid instrument for measuring self-criticism in clinical patients is needed. One instrument that is recently used more frequent is the Forms of Self-Criticizing / -attacking and Self-Reassuring Scale (FSCRS), developed by Gilbert et al. (2004) in order to investigate of the role of self-criticism in types of psychopathology. The instrument was developed and

validated in the context of depression treatment. Based on observed thoughts during clinical sessions a set of items was derived. These items were evaluated in a sample of female students and their statements about how they tend to handle situations of failure, which resulted in a 22- item questionnaire. The questionnaire assesses three underlying factors that were found with exploratory factor analysis. Two factors incorporate the concept of self-criticism: *inadequate self* (9 items), capturing inadequate feelings that one deserves criticism and dwelling on failure and *hated self* (5 items), a tendency to be more disgusted and aggressive with oneself, and one factor for *reassured self* (7 items). Self-reassuring within this instrument is described as encouraging and warm disposition towards oneself in terms of failure.

Confirmatory factorial analyses in both non-clinical (Kupeli, Chilcot, Schmidt, Campbell, & Troop, 2013) and clinical samples (Baião, Gilbert, McEwan, & Carvalho, 2015) showed a good fit for the assumed three-factor (inadequate self, hated self and reassured self) versus two factor (self-criticism and self-reassuring) model structure. Furthermore, the three-factor model has been confirmed in a recently developed short-form of the FSCRS in Dutch language in a non-clinical sample (Sommers-Spijkerman et al., 2017).

When constructing and evaluating the FSCRS, Gilbert and his colleagues (2004) found good levels of internal consistency for all three subscales: Cronbach's alpha ranged from .89 to .91 for inadequate-self, .82 to .89 for hated-self, and .82 to .88 for reassured-self in non-clinical samples whereas Cronbach's alphas in clinical samples ranged from .87 to .89 for inadequate-self, .83 to .86 for hated-self, and .85 to .87 for reassured-self (Sommers-Spijkerman et al., 2017; Baião et al., 2015; Kupeli et al., 2013).

Moreover, the FSCRS has been translated and validated for use in various countries, such as Portugal (Palmeira et al., 2017), Sweden (Lekberg, & Wester, 2012), Czech Republic

(Halamova, Kanovsky, & Pacuchova, 2017) and the Netherlands (Sommers-Spijkerman et al., 2017).

Although the Dutch (short form) version of the FSCRS has been validated within a non-clinical context as well as in a sample with low to moderate levels of mental wellbeing, its psychometric properties within specified psychopathological populations, such as within eating disorder patients, are yet to be investigated. As self-criticism is an important factor in the prediction of therapeutic outcomes, especially within eating disorder patients, an instrument such as the FSCRS and its validity and reliability need to be investigated in these populations as well. The present study will therefore investigate the psychometric properties of the Dutch version of the FSCRS within a sample of patients diagnosed with eating disorders. Specifically, the factorial validity, internal consistency and convergent validity of the instrument will be investigated.

2 Method

Participants and procedure

A cross-sectional survey-based research design employing purposive sampling was used to obtain data for this project. The participants (N= 502) for the survey were recruited at the Human Concern Foundation Center for Eating Disorders in Amsterdam, where specialist treatment is offered for men and women diagnosed with eating disorders. The treatment has several contra-indications: patients are excluded when they (1) have any main diagnosis other than eating disorder, (2) show actual signs of suicidal actions, (3) suffer from severe somatic parameters, (3) have a Body-Mass Index (BMI) lower than 13.

For obtaining information about the individual's diagnosis, the Diagnostic and Statistical Manual of Mental Disorders – Fifth Edition (DSM-V; American Psychiatric Association, 2013) and the Eating Disorder Examination Questionnaire (EDE-Q; Fairburn & Beglin, 1994; translated by Van Furth, 2000) are used. After stating a diagnosis, the patients process the Routine Outcome Monitoring, where patients' individual parameters are evaluated during the progress of treatment. The present study only included the surveys that were completed during the intake before treatment. Patients signed informed consent for the use of anonymized surveys for research purposes. Ethical approval from commission of ethical psychology of the University of Twente was obtained to use the data.

Instruments

The following measuring instruments were used to obtain data for this study:

A *demographic questionnaire* was used to gather information regarding participants gender and age.

The Dutch version of the *Forms of Self-Criticizing/ -Attacking and Self-Reassuring Scale* (FSCRS), developed by Gilbert et al. (2004) and translated by Sommers-Spijkerman et. al (2017) was used to measure self-criticism and its underlying components. The self-rated

instrument consisted out of 22 items which are rated on a 5-point Likert type scale ranging from 1 (‘Not at all like me’) to 5 (‘Extremely like me’). These items were grouped into three components of self-criticism: the inadequate self (‘I am easily disappointed with myself’), the hated self (‘I have become so angry with myself that I want to hurt or injure myself’) and the reassured self (‘I am gentle and supportive with myself’). The instrument showed acceptable levels of internal consistency in other studies with Cronbach’s alpha coefficients ranging .90 for inadequate self, .83 for hated self, and .88 for reassured self (Kupeli, et al., 2013).

The Dutch version of the *Mental Health Continuum Short Form* (MHC-SF; Keyes, 2002), translated by Lamers, Westerhof, Bohlmeijer, ten Klooster, and Keyes (2011), was used to measure mental wellbeing within this sample. On a 14-item questionnaire, the participants state how often they experienced varied factors of mental wellbeing during the past period. Mental wellbeing is captured in three dimensions: emotional wellbeing (‘…happy?’), social wellbeing (‘…that the way our society works makes sense to you?’) and psychological wellbeing (‘that you had experiences that challenged you to grow and become a better person?). On a 6-point scale, the answers range from 0 (‘Never’) to 5 (‘Every day’). The Dutch version of the MHC-SF is well validated and shows good reliability (Cronbach’s alpha) for the three dimensions, being .74 for social wellbeing and .83 for emotional and psychological wellbeing respectively (Lamers, et al. 2011). For the present study, similar alpha levels of .86, .72 and .83 were found for the subscales respectively.

Statistical Analyses

For the statistical analyses, the present study employed both SPSS version 24 (IBM, 2016) as well as Mplus Version 8 (Muthén & Muthén, 2017).

First, descriptive statistics were calculated to provide *demographic characteristics* of the sample. Crosstabulations were made to provide information about gender and age within

the groups. A MANOVA analysis was run to examine possible differences in scores on the FSCRS subscales between the groups which could influence the psychometric properties of the instrument in the total sample.

For *factorial validity*, a confirmatory factor analytical approach utilizing structural equation modelling (SEM) with the robust maximum likelihood estimator (MLR) was employed. This estimator assumes normality in missing data and is robust to non-normality and non-independence of observations (Muthén & Muthén, 2017). For evaluating the proposed fit of the data against the competing models, the following fit indices were utilized to evaluate the fit: Chi-square (χ^2) statistics, Comparative Fit Index (CFI), Tucker Lewis Index (TLI), Root Mean Square Error of Approximation (RMSEA) and Standardized Root Mean Square Residual (SRMR). In table 2 the corresponding cut-off scores are displayed, as indicated by Hu and Bentler (1999). Additionally, the Akaike Information Criterion (AIC) and Bayesian Information Criterion (BIC) were used to evaluate the comparing fit of the competing models. Finally, differences in fit were statistically tested using Satorra-Bentler scaled difference χ^2 ($\Delta SB\chi^2$) tests (Satorra & Bentler, 2010).

Table 1 *Fit indices: Indications for acceptable - and good fit indices*

Fit Indices	Acceptable Fit	Good Fit
Chi-Square ($\chi 2$)	Lowest value in comparative measurement models	
Root Mean Square Error of Approximation (RMSEA)	< .08	< .06
Standardized Root Mean Square Residual (SRMR)	< .10	< .08
Comparative Fit Index (CFI)	> .90	> .95 and < 0.99
Tucker Lewis Index (TLI)	> .90	> .95 and < 0.99
Akaike Information Criterion (AIC)	Lowest value in comparative measurement models	
Bayesian Information Criterion (BIC)	Lowest value in comparative measurement models	

Internal consistency was computed with the Cronbach's alpha, where a value of ≥ .70 is considered fair to good whereas values > .90 are considered excellent (Cicchetti, 1994).

For *convergent validity* measurement, Pearson correlation analyses were conducted. Correlations between .10 and .30 were considered small, correlations between .30 and .50 were considered moderate and correlations > .50 were considered strong (Cohen, 1988). It was hypothesized that a valid measure of self-criticism should have a negative relation to mental wellbeing. Specifically, the three subscales of the FSCRS were expected to be differentially correlated with the different scales of mental wellbeing measured with MHC-SF. The two scales of self-criticism (i.e. inadequate self and hated self) were expected to have negative correlations with the MHC-SF scale of emotional wellbeing. Feelings of inadequacy, internal put-downs and feeling subordinate (Gilbert et al., 2004, 2006) are seen as conflicting factors to being happy, interested in and satisfied with life (Keyes, 2002). Therefore, a moderate to strong negative correlation between inadequate self and emotional wellbeing as well as between hated self and emotional wellbeing was expected. Additionally, both inadequate self and hated self were expected to have a strong negative correlation with the scale of psychological wellbeing. Feelings of self-hating and self-attacking, that consist of wanting to harm and destroy the self (Gilbert et al., 2004) are conceptionally negatively related to self-acceptance, environmental mastery, positive relations with others, personal growth, autonomy, and purpose in life (Ryff, 1989). At last, inadequate self and hated self were expected to have weak to moderate negative correlations with social wellbeing. As the concept of self-criticism is defined as self-to-self relating (Kupeli et al., 2013; Gilbert et al., 2004) the connection to social wellbeing (self-to-others relationship) is expected to be lower than with the intra-individual constructs of emotional and psychological wellbeing.

The concept of self-reassurance is considered as part of self-compassion and self-acceptance (Palmeira et al., 2017). As self-acceptance is defined as a positive attitude towards oneself (Ryff, 1989) it can be seen as an element of psychological wellbeing (Ryff, 1989), therefore, the scale of reassured self was expected to have a strong positive correlation with the MHC-SF scale psychological wellbeing. Also reassured self was expected to have at least moderate correlations with the other scales of the MHC-SF (emotional and social wellbeing) as 'self-reassurance contributes to mental health and wellbeing…' (p. 2, Sommers-Spijkerman et al., 2017; Gilbert et al., 2004).

3 Results

Table 2 displays the demographic characteristics of the sample (N=502), stratified by the diagnosed groups. The majority of the sample was female with a diagnosis of AN. Furthermore, most were younger than 30 years old. There was no statistically significant difference between group means for any of the FSCRS subscale scores as determined by MANOVA, see Table 3.

Table 2 *Demographic characteristics per Diagnosis (N=502)*

Variable	AN[a] (n= 177)	BN (n= 114)	BED (n=65)	OSFED (n=146)	Total
Gender, n (%)					
Female	174 (98.9%)	114 (100%)	64 (98.5%)	144 (98.6%)	496 (99.0%)
Male	2 (1.1%)	0 (0%)	1 (1.5%)	2 (1.4%)	5 (1.0%)
Age in yrs, n (%)					
16 to 20	72 (40.7%)	20 (17.5%)	7 (10.8%)	35 (24.0%)	134 (26.7%)
21 to 30	68 (38.4%)	63 (55.3%)	28 (43.1%)	77 (52.7%)	236 (47.0%)
31 to 40	23 (13.0%)	17 (14.9%)	21 (31.3%)	20 (13.7%)	81 (16.1%)
41 to 50	11 (6.2%)	10 (8.8%)	4 6.2%)	7 (4.8%)	32 (6.4%)
51 +	3 (1.7%)	4 (3.5%)	5 (7.7%)	7 (4.8%)	19 (3.8%)

[a] = one missing value in Gender; [b]= non-significant difference between groups ($p>0.05$); AN, Anorexia Nervosa; BN, Bulimia Nervosa; BED, Binge Eating Disorder; OSFED, Other Specified Feeding -/ Eating Disorder

Table 3 *MANOVA Differences of FSCRS Subscale Scores in Eating Disorders*

Variable	AN	BN	BED	OSFED	*p* Value[a]	Partial η^2
	M (SD)	M (SD)	M (SD)	M (SD)		
IS	2.78 (0.73)	2.87 (0.78)	2.76 (0.73)	2.77 (0.71)	0.64	0.00
RS	1.32 (0.78)	1.51 (0.73)	1.39 (0.78)	1.40 (0.77)	0.26	0.01
HS	1.90 (0.91)	1.92 (0.84)	1.73 (0.81)	1.87 (0.83)	0.54	0.00

[a]= significance at .05 level; IS, Inadequate Self; RS, Reassured Self; HS, Hated Self; AN, Anorexia Nervosa; BN, Bulimia Nervosa; BED, Binge Eating Disorder; OSFED, Other Specified Feeding -/ Eating Disorder; M, Mean; SD, Standard deviation

Analysis of Wilk's Lambda value showed no statistically significance ($p > .05$) therefore no significant differences in any of the subscale scores between the groups of eating disorder were found; all groups of eating disorders experienced similar levels of self-criticism.

Factorial validity (Confirmatory Factor Analysis)

To determine the factor structure of the FSCRS, confirmatory factor analyses were used comparing the proposed three-, two-, and one factor models for self-criticism in the total sample. A competing measurement model strategy was adopted to systematically evaluate and compare the three proposed measurement models. All 22 items were utilized in each analysis.

The following models were compared within the confirmatory factor analysis: firstly, a three-factor model (Model A) is measured, comprised out of the previously established constructs of inadequate self, hated self and reassured self. Second a two-factor model (Model B), assumed to be comprised out of two components, one component for self-criticism (comprised out of the inadequate self and hated self) and self-reassurance respectively was specified, as also suggested by Kupeli et al. (2013). Then, a strict unidimensional model was tested (Model C), assuming that all items load directly on a single

construct of self-criticism. Correlations between factors in all models was permitted, no item error terms correlation was allowed. Each item was permitted to load on one factor only.

Table 3 *Fit statistics for competing measurement models*

Model	χ^2	*df*	TLI	CFI	RMSEA	SRMR	AIC	BIC
Model A	622.16	206	0.89	0.90	0.06	0.05	30010.07	30301.16
Model B	784.88	208	0.84	0.86	0.07	0.06	30184.26	30466.91
Model C	1049.43	209	0.77	0.79	0.09	0.07	30476.93	30755.36
Model D	574.94	204	0.90	0.92	0.06	0.05	29930.05	30229.57

Model A= three factor model; Model B= two factor model; Model C= one factor model; Model D= three factor model with 2 residual variance correlations allowed, with items 1 and 2; 2 and 6; χ^2, Chi-square; *df*, degrees of freedom; TLI, Tucker-Lewis Index; CFI, Comparative Fit Index; RMSEA, Root Mean Square Error of Approximation; SRMR, Standardised Root Mean Square Residual; AIC, Akaike Information Criterion; BIC, Bayes Information Criterion

Table 3 indicates that a three-factor model fitted the data better than the one and two factor models. χ2, AIC and BIC, were lower for Model A than for Model B. Although most fit indices of the three-factor model showed an acceptable and good fit (see Table 1) the TLI value was just below the cut-off for acceptable fit. Therefore, a fourth model (Model D) was estimated to see which conditions would increase model fit. In order to get a sufficiently fitting model, according to the pre-defined cut-off values, two sets of two items' residual variances were allowed to correlate within the three-factor model. Modification indices were used to estimate the proposed impact of residual variance correlations between items. With permitting correlations of residual variances for items 1 ('I am easily disappointed with myself') and 2 ('There is a part of me that puts me down'), as well as item 2 and 6 ('There is a part of me that feels I am not good enough') being correlated, the model fitted better than when no covariation is permitted (χ^2 = 574.94; *df*=204; TLI= 0.90; CFI= 0.92; RMSEA= 0.06; SRMR= 0.05; AIC= 29930.05; BIC= 30229.57). All indices, including the TLI value,

showed acceptable/good model fit after allowing residual variances to correlate, no further correlations are needed.

Additionally, Satorra-Bentler scaled χ^2 was calculated (Satorra & Bentler, 2010) to compare the statistically significant differences in χ^2 between the models. These differences are shown in Table 4, supporting the finding that the three-factor model with parcelled items (Model D) provided statistically significant best model fit (see Model A vs Model D).

Table 4 *Model Comparison with Differences in Satorra-Bentler* χ^2

	C1	**C0**	**D1**	**D0**	**F1**	**F0**	**ΔSBχ^2**	***df***	***p* value**
Model A vs Model B	1.10	1.09	206.00	208.00	622.16	784.88	302.86	2	<0.01
Model A vs Model C	1.10	1.10	206.00	209.00	622.16	1049.43	407.32	3	<0.01
Model A vs Model D	1.10	1.08	206.00	204.00	622.16	547.94	-38.03	2	<0.01
Model B vs Model C	1.09	1.10	208.00	209.00	784.88	1049.43	127.95	1	<0.01
Model B vs Model D	1.09	1.08	208.00	204.00	784.88	547.94	-181.76	4	<0.01
Model C vs Model D	1.10	1.08	209.00	204.00	1049.43	547.94	-344.87	5	<0.01

C1= scaling correction for compared model; C0= scaling correction for nested model; D1= degrees of freedom in compared model; D0= degrees of freedom in nested model; F1= χ^2 of compared model; F0= χ^2 of nested model; ΔSBχ^2= Difference in Satorra-Bentler scaled χ^2; calculated with ΔSBχ^2= (d0 * c0 - d1*c1)/ (d0 - d1)

The items loadings for the three variables as found in Model D are displayed in Table 4, which shows adequate loadings for all items on their latent factors (>.35; Wang & Wang, 2012). Item loadings for inadequate self ranged from .39 to .73. On the variable of reassured

self the loadings ranged from .40 to .80 whereas a ranging from .67 to .77 for the items on hated self component is observed.

Table 4 *Standardized factor loadings for latent variables*

Factor	Item No	Item text	Loading	S.E.
Inadequate Self	1	I am easily disappointed with myself	0.69*	0.03
	2	There is a part of me that puts me down	0.73*	0.03
	4	I find it difficult to control my anger and frustration at myself	0.46*	0.04
	6	There is a part of me that feels I am not good enough	0.66*	0.04
	7	I feel beaten down by my own self-critical thoughts	0.70*	0.03
	14	I remember and dwell on my failings	0.57*	0.04
	17	I can't accept failures and setbacks without feeling inadequate	0.54*	0.04
	18	I think I deserve my self-criticism	0.69*	0.03
	20	There is a part of me that wants to get rid of the bits I don' like	0.39*	0.05
Reassured Self	3	I am able to remind myself of positive things about myself	0.67*	0.03
	5	I find it easy to forgive myself	0.56*	0.04
	8	I still like being me	0.80*	0.03
	11	I can still feel lovable and acceptable	0.66*	0.03
	13	I find it easy to like myself	0.73*	0.03
	16	I am gentle and supportive with myself	0.62*	0.04
	19	I am able to care and look after myself	0.40*	0.04
	21	I encourage myself for the future	0.58*	0.04
Hated Self	9	I have become so angry with myself that I want to hurt or injure myself	0.67*	0.03
	10	I have a sense of disgust with myself	0.76*	0.02
	12	I stop caring about myself	0.73*	0.02
	15	I call myself names	0.68*	0.03
	22	I don't like being me	0.77*	0.02

* p <0.001; No cross-loading items; S.E. = Standard Error

Internal Consistency

Table 5 provides an overview of the descriptive statistics (means, standard deviations, skewness and kurtosis), as well as the Cronbach Alphas of the two instruments used in this study. Specifically, it indicates good internal consistency for all subscales of the FSCRS. The levels of alpha for the subscales within the total sample were .84 for inadequate self, .85 for hated self and .83 for reassured self, respectively. Within the diverse types of eating disorders, alpha did not substantially differ with alpha's ranging from .82 and .84 for the scale of inadequate self, ranging from .82 to .88 for the scale of hated self and alpha ranging from .80 to .84 for the scale of reassured self (Table 5).

Table 5 *Descriptive statistics, alpha coefficients and Pearson correlations*

Variable	**x̄**	***σ***	**Skew**	**Kurt**	**α**	**AN α**	**BN α**	**BED α**	**OSFED α**	**1**	**2**	**3**
Self-Criticism												
1 IS	2.79	0.74	-.69	-.01	.84	.84	.83	.83	.82	-	-	-
2 HS	1.87	0.86	.26	-.84	.85	.88	.84	.82	.83	.69	-	-
3 RS	1.39	0.76	.45	-.21	.83	.84	.80	.84	.84	-.56	-.61	-
Mental Health												
EWB	2.34	1.17	-.07	-.76	.86	.88	.86	.83	.86	-.45	-.51	.61
SWB	2.03	0.99	.37	-.28	.72	.68	.76	.73	.73	-.42	-.38	.49
PWB	2.36	0.74	.18	-.48	.83	.80	.84	.87	.85	-.51	-.53	.69

IS, Inadequate Self, HS, Hated Self, RS, Reassured Self; EWB, Emotional Wellbeing; SWB, Social Wellbeing; PWB, Psychological Wellbeing; x̄, Mean; σ, Standard Deviation; Skew, Skewness; Kurt, Kurtosis; α, Cronbach's alpha; AN, Anorexia Nervosa; BN, Bulimia Nervosa; BED, Binge Eating Disorder; OSFED, Other Specified Feeding-/Eating Disorder

Convergent Validity

All correlations between the scales of self-criticism and mental wellbeing were >.30. Inadequate self had, as hypothesized, a moderate negative correlation with emotional wellbeing (r= -.45) and social wellbeing (r= -.42) respectively. As expected, a strong negative relationship was found with psychological wellbeing (r= -.51). The scale of hated self had a strong negative correlation with emotional wellbeing (r= -.51) and psychological wellbeing (r= -.53). A moderate negative relation was found with social wellbeing (r= -.38) just as expected. Reassured self had strong positive correlations both with emotional wellbeing (r= .61) and psychological wellbeing (r= .69) and a moderate correlation with social wellbeing (r= .49). The observed pattern of the correlations with aspects of mental health was exactly as expected. All the observed correlation coefficients were in accordance with the a-priori hypotheses.

4 Discussion

The present study evaluated the psychometric properties of the FSCRS in a sample of patients with eating disorders in the Netherlands. The factorial validity, internal consistency and convergent validity of the instrument were investigated. With the findings on these analyses, the present study provides confirmation that the FSCRS is a reliable and valid instrument to assess features of self-criticism and self-reassurance within Dutch patients with eating disorders.

Analysis of the scores on the FSCRS showed that there is no significant difference in the scores on the instrument between the groups of eating disorders, supporting the finding of Duarte, Ferreira and Pinto-Gouveia (2016), who stated that all types of eating disorders experience the same level of self-criticism and self-reassurance. Therefore, all analyses on validity and reliability were conducted for the entire sample and for the separate diagnoses of eating disorders.

Confirmatory factor analysis was established to investigate whether the expected three-factor model can be confirmed. The model as assumed by Gilbert et al. (2004) and later confirmed by various researchers (Sommers-Spijkerman et al., 2017; Baião et al., 2015; Kupeli et al., 2013) was found to have acceptable fit in the present sample of Dutch patients with eating disorders. Although a three-factor model, consisting of the components inadequate self, hated self and reassured self, fits the data better than other models, not all fit indices for the three-factor model were sufficient. Therefore, residual variances of two pairs of items were allowed to correlate to improve model fit, just so the fit indices criteria are met on an acceptable level. Justification of residual variance correlation has been discussed in research, nevertheless it has been also used for validation of the FSCRS. Baião et al. (2015) confirmed the factor structure only with allowing several error correlations (5 item pairs in total). Another validation study (Kupeli et al., 2013) allowed items to double-load or even

removed one item from the analysis. In the current study the affected items load on the same factor (inadequate self; Gilbert et al., 2004). Anyway, these error correlations indicate that there is any other influence than the shared factor that displays a covariance within the items. Whereas Baião and colleagues (2015) and Kupeli and colleagues (2013) do not give any explanation for adjusting the items, the association of items within the present study may depend on unsatisfactory translation of the questionnaire. An additional exploratory factor analysis for the Dutch version of the FSCRS could provide information about any association due to the use of similar words within the items.

The fact that the model fit indices were not found at a sufficient level within a three-factor model, without any correlation permitted, indicates that there is a need to follow up on the factor structure of the FSCRS. Independent studies in new samples of clinical and non-clinical populations can give further information and about the replicability of error and residual variance correlations.

The present study found good internal consistency (Cronbach's alpha) for all subscales of the FSCRS for all groups of eating disorders. Cronbach's alpha above .70, such as found in the present study, are considered as fair to good (Cicchetti, 1994). Here, the alphas were similar for all subscales (all >.80), confirming the findings of other validation studies (Baião et al., 2015; Kupeli et al., 2013). The Cronbach's alpha that are found are sufficient to compare the reliability in groups. However, in this context the FSCRS was used to monitor the patient's progress in treatment. For this purpose, the reliability is not sufficient. In order to confirm the reliability as found in the present study, test-retest reliability of the scores should be examined in future studies.

The correlations that were found between the FSCRS and MHC-SF confirmed that the FSCRS has good convergent validity. It was predicted that a negative relationship should exist between the aspects of self-criticism and mental wellbeing, whereas self-reassurance

should correlate positively with mental wellbeing. The strength of the correlations was found as expected. For further confirmation of the FSCRS' validity, research could investigate in how far correlations with other (valid) instruments for self-criticism are found (concurrent validity), such as the Levels of Self-Criticism Scale (LOSC; Thompson & Zuroff, 2000). Furthermore, incremental validity should be examined to investigate whether the FSRCS adds value in predicting self-criticism or if other instruments sufficiently represent the measurement of symptoms of self-criticism in practice.

Some limitations concerning the sample are noted for the present study. The sample of eating disorders does not include patients with other main diagnosis. Patients who are diagnosed with eating disorders as second diagnosis were excluded from the treatment. Research shows that eating disorders are highly comorbid with at least one more diagnosis and around 97% of the patients suffer from a second disorder (Blinder, Cumella, & Sanathara, 2006) whereby it is not clear which diagnosis requires more attention and displays more severity. It can be concluded that comorbidity might have influence on the quality of the instrument, when patients suffer from other psychopathologies as well. Generalizability of the finding for comorbid patients is therefore questioned and should be addressed in further clinical research.

Moreover, samples of all validation studies (Sommers-Spijkerman et al., 2017; Baião et al., 2015; Kupeli et al., 2013) contained mainly females. No studies stated gender differences in self-criticism in clinical samples, so it cannot be concluded whether men experience self-criticism conceptually different than women. The present study is also mostly limited to females rather than males but as eating disorders are mostly prevalent in women (Hudson et al., 2007), it is concluded that the findings of the present study can be generalized for patients with eating disorders. Since Baião et al. (2015) found that in non-clinical samples the males were significantly more self-reassured and less self-critical than females, the FSCRS cannot

give a generalization if the instrument is valid and reliable for males in non-clinical samples. Gender differences in the concept of self-criticism in both clinical and non-clinical samples could also be addressed with measurement invariance testing. It is advised to prospectively provide genderwise norms for the FSCRS scores.

Although scores on the FCSRS showed that there is no statistically significant difference in the level of self-criticism between the different types of eating disorders, further research on measurement invariance between the groups can provide more detailed information about factorial validity for the groups respectively. The current study is limited to the total group of eating disorder patients as sample size did not allow separate measurement invariance analysis for AN, BN BED and OSFED respectively.

As stated before, another limitation concerns the translation of the instrument. The Dutch version of the FSCRS was not translated within an expert panel. The translation was conducted as word-to-word translation, which may not adequately capture linguistic and/or cultural differences. Measurement invariance analysis of the translated versions can show whether there are conceptual differences of self-criticism within diverse cultures. The equivalency of words within the translated version is not checked, therefore the Dutch version contains some risk for construct, method and item bias (Van de Vijver & Hambleton, 1996). All these biases also can decrease validity of an instrument (Sperber, 2004; Van de Vijver & Hambleton, 1996). Sperber (2004) states multiple possibilities of adequate development of instruments for cross-cultural usage, such as back-translation technique and two-source translation, that decrease the possibility of biases occurring.

Although validity and reliability are confirmed in the present study, further examinations such as incremental and concurrent validity, measurement invariance testing and reliability testing are needed to provide more evidence for the psychometric properties of the FSCRS.

In summary, the FSCRS so far has been found as valid and reliable instrument to measure self-criticism and self-reassurance in Dutch patients with eating disorders. The FSCRS can help clinicians and researchers to make better conclusions and interpretations of the scores on the Forms of Self-Criticizing /- Attacking and Self-Reassuring Scale. For practice use, the FSCRS needs some improvement in order to monitor changes on self-criticism during treatment.

References

American Psychiatric Association. (2013). *Diagnostic and Statistical Manual of Mental Disorders* (5th ed.). Arlington, VA: American Psychiatric Publishing

Baião, R., Gilbert, P., McEwan, K., & Carvalho, S. (2015). Forms of Self- Criticizing and Attacking and Self-Reassuring Scale: Psychometric properties and normative study. *Psychology and Psychotherapy: Theory, Research and Practice*, *88*(4), 438-452. doi: 10.1111/papt.12049

Blinder, B. J., Cumella, E. J., & Sanathara, V. A. (2006). Psychiatric comorbidities of female inpatients with eating disorders. *Psychosomatic Medicine*, *68*(3), 454-462. doi: 10.1097/01.psy.0000221254.77675.f5

Cicchetti, D. V. (1994). Guidelines, criteria and rules of thumb for evaluating normed and standardized assessment instruments in psychology. Psychological Assessment, 6, 284–290. doi: 10.1037/10403590.6.4.284

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences (2nd ed.).* Hillsdale, NJ: Erlbaum

Cole, D. A., Ciesla, J. A., & Steiger, J. H. (2007). The insidious effects of failing to include design-driven correlated residuals in latent-variable covariance structure analysis. *Psychological Methods*, *12*(4), 381.doi: 10.1037/1082-989X.12.4.381

Duarte, C., Ferreira, C., & Pinto-Gouveia, J. (2016). At the core of eating disorders: Overvaluation, social rank, self-criticism and shame in anorexia, bulimia and binge eating disorder. *Comprehensive psychiatry*, *66*, 123-131. doi: 10.1016/j.comppsych.2016.01.003

Dunkley, D. M., & Grilo, C. M. (2007). Self-criticism, low self-esteem, depressive symptoms, and over-evaluation of shape and weight in binge eating disorder patients. *Behaviour Research and Therapy, 45*(1), 139-149. doi: 10.1016/j.brat.2006.01.017

Fairburn, C.G. & Beglin, S.J. (1994). Assessment of Eating Disorders: Interview or Self-Report Questionnaire? International Journal of Eating Disorders, 16, 363 – 370. doi: 10.1002/1098-108X(199412)16

Fairburn, C. G., & Harrison, P. J. (2003). Eating disorders. *The Lancet*, *361*(9355), 407-416. doi: 10.1016/S0140-6736(03)12378-1

Falconer, C. J., King, J. A., & Brewin, C. R. (2015). Demonstrating mood repair with a situation - based measure of self - compassion and self - criticism. *Psychology and Psychotherapy: Theory, Research and Practice*, *88*(4), 351-365. doi: 10.1111/papt.12056

Fennig, S., Hadas, A., Itzhaky, L., Roe, D., Apter, A., & Shahar, G. (2008). Self - criticism is a key predictor of eating disorder dimensions among inpatient adolescent females. *International Journal of Eating Disorders*, *41*(8), 762-765. doi: 10.1002/eat.20573

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., 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., & Procter, S. (2006). Compassionate mind training for people with high shame and self - criticism: Overview and pilot study of a group therapy approach. *Clinical Psychology & Psychotherapy*, *13*(6), 353-379. doi: 10.1002/cpp.507

Halamova, J., Kanovsky, M., & Pacuchova, M. (2017). Robust psychometric analysis and factor structure of the Forms of Self-criticizing/attacking and Self-reassuring scale. Ceskoslovenska Psychologie, 61(4), 331-349.

Hay, P., Girosi, F., & Mond, J. (2015). Prevalence and sociodemographic correlates of DSM-5 eating disorders in the Australian population. *Journal of Eating Disorders, 3, 19.* doi: 10.1186/s40337-015-0056-0

Hu, L., & Bentler, P. M. (1999). Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. *Structural Equation Modeling, 6(1),* 1-55. doi: 10.1080/10705519909540118

Hudson, J. I., Hiripi, E., Pope, H. G., & Kessler, R. C. (2007). The prevalence and correlates of eating disorders in the national comorbidity survey replication. *Biological Psychiatry*, *61*(3), 348–358. doi: 10.1016/j.biopsych.2006.03.040

IBM SPSS Statistics. (2016). *SPSS for Windows 24*. Chicago: SPSS Inc.

Kannan, D., & Levitt, H. M. (2013). A review of client self-criticism in psychotherapy. *Journal of Psychotherapy Integration*, *23*(2), 166. doi: 10.1037/a0032355

Keyes, C. L. (2002). The mental health continuum: From languishing to flourishing in life. *Journal of Health and Social Behavior*, 207-222. Retrieved from http://www.jstor.org/stable/3090197

Kupeli, N., Chilcot, J., Schmidt, U. H., Campbell, I. C., & Troop, N. A. (2013). A confirmatory factor analysis and validation of the forms of self-criticism/reassurance

scale. *British Journal of Clinical Psychology, 52(1),* 12-25. doi:10.1111/j.2044-8260.2012.02042.x

Lamers, S.M.A., Westerhof, G.J., Bohlmeijer, E.T., ten Klooster, P.M., & Keyes, C.L.M. (2011). Evaluating the psychometric properties of the Mental Health Continuum-Short Form (MHCSF). *Journal of Clinical Psychology*, 67*(1),* 99-110. doi: 10.1002/jclp.20741

Lekberg, A., & Wester, S. (2012). *Självkritik eller självuppmuntran? Hur vi beter oss mot oss själva vid motgångar: En utprövning av instrumenten FSCRS och FSCS*. (Masters dissertation) Retrieved from http://lup.lub.lu.se/student-papers/record/2827165

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

Muthén, L. K., & Muthén, B. O. (2017). *Mplus User's Guide (Eight Edition)*. Los Angeles, CA: Muthén & Muthén.

Palmeira, L., Pinto-Gouveia, J., Cunha, M., & Carvalho, S. (2017). Finding the link between internalized weight-stigma and binge eating behaviors in Portuguese adult women with overweight and obesity: The mediator role of self-criticism and self-reassurance. *Eating Behaviors, 26*, 50-54. doi: 10.1016/j.eatbeh.2017.01.006

Pinto - Gouveia, J., Ferreira, C., & Duarte, C. (2014). Thinness in the pursuit for social safeness: An integrative model of social rank mentality to explain eating psychopathology. *Clinical Psychology & Psychotherapy, 21*(2), 154-165. doi: 10.1002/cpp.1820

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*(5), 571-584. doi: 10.1023/A:1005566112869

Ryff C. D. (1989). Happiness is everything, or is it? Explorations on the meaning of psychological well-being. *Journal of Personality and Social Psychology*, 57, 1069–1081. doi: 10.1037/0022-3514.57.6.1069

Satorra, A., & Bentler, P.M. (2010). Ensuring positiveness of the scaled difference chi-square test statistic. *Psychometrika*, *75*, 243-248. doi: 10.1007/s11336-009-9135-y

Smink, F. R., Van Hoeken, D., & Hoek, H. W. (2012). Epidemiology of eating disorders: incidence, prevalence and mortality rates. *Current Psychiatry Reports*, *14*(4), 406-414. doi: 10.1007/s11920-012-0282-y

Sperber, A. D. (2004). Translation and validation of study instruments for cross-cultural research. *Gastroenterology*, *126*, 124-128. doi: 10.1053/j.gastro.2003.10.016

Thew, G. R., Gregory, J. D., Roberts, K., & Rimes, K. A. (2017). The phenomenology of self-critical thinking in people with depression, eating disorders, and in healthy individuals. *Psychology and Psychotherapy*, doi:10.1111/papt.12137

Thompson, R., & Zuroff, D.C. (2000). *The Levels of Self-Criticism Scale*. Poster session: APA: Assessment and Diagnosis, Washington Convention Centre.

Van Furth, E. F. (2000). *Nederlandse vertaling van de EDE-Q.* Leidschendam: Robert-Fleury Stichting.

Van de Vijver, F., & Hambleton, R. K. (1996). Translating tests. *European Psychologist*, 1(2), 89-99. doi: 10.1027/1016-9040.1.2.89

Wang, J., & Wang, X. (2012). *Structural equation modelling: Applications using Mplus.* Chichester, UK: Wiley & Sons.