The Development of the 'Forms of Responding to Self-Critical Thoughts Scale' (FoReST)

Ross G. White[1], Peter Larkin[2], Judith McCluskey[2], Joda Lloyd[3], and Hamish J. McLeod[4]

[1]Institute of Psychology, Health and Society
University of Liverpool,
G.10, Ground floor, Whelan Building Quadrangle
Brownlow Hill
Liverpool
L69 3GB
U.K.
*Corresponding author
Email: ross.white@liverpool.ac.uk

[2]NHS Greater Glasgow and Clyde
1055 Great Western Rd,
Glasgow
G12 0XH
U.K.

[3]Institute of Management Studies
Goldsmiths, University of London
New Cross
London
SE14 6NW
U.K.

[4]Institute of Health and Wellbeing
University of Glasgow
Gartnavel Royal Hospital
1055 Great Western Road
Glasgow
G12 0XH
UK

Highlights

- Self-critical thoughts are associated with a range of mental health difficulties.
- High levels of psychological inflexibility are associated with increased distress.
- A measure of psychological inflexibility related to self-critical thoughts is needed.
- The factor structure and validity of a new measure called the FoReST are explored.
- The FoReST is a valid measure of inflexible responding to self-critical thoughts.

Abstract

Background: Self–critical thoughts are a feature of many mental health problems. Adopting a psychologically flexible response to thoughts has been highlighted as a key determinant of wellbeing. But, the measurement of psychological flexibility (PF) in relation to self-critical thoughts is under-developed. This paper reports on the development of the *Forms of Responding to Self-Critical Thoughts Scale* (FoReST).

Method: Study One involved the development and exploratory factor analysis of the FoReST in a convenience sample of 253 non-clinical adults. Study Two was a confirmatory factor analysis study of the FoReST in a sample of 110 University students. Study 3 explored the convergent and concurrent validity of the FoReST by examining associations with measures of similar constructs (PF, self-compassion, self-criticism) and relevant mental health measures (anxiety, depression, distress). Study 3 also explored the FoReST's incremental validity for predicting depression and anxiety levels beyond an established measure of self-critical thinking.

Results: In Study One, exploratory factor analysis produced a 9-item/2-factor solution (*unworkable action* and *mindful acceptance*). The measure demonstrated good internal consistency. In Study 2, the confirmatory factor analysis also indicated a 2-factor model (*unworkable action* and *mindful acceptance*) and overall internal consistency that was excellent. In Study Three, the FoReST demonstrated high convergent validity with similar measures, and good concurrent validity with mental health outcomes. Analyses also indicted good incremental validity for the FoReST for predicting HADS depression and anxiety scores.

Conclusions: The FoReST appears to be a psychometrically sound measure suitable for measuring change processes in third wave therapies such as Acceptance and Commitment Therapy, Mindfulness-based Interventions and Compassion Focused therapy.

Keywords

Self-criticism; Depression; Psychological flexibility; Acceptance and Commitment Therapy; Mindfulness; Compassion Focused Therapy

Introduction

Excess self-criticism can be understood as a form of stressful self-harassment, which undermines healthy self-acceptance (Gilbert, 2004). Elevated levels of self-critical thoughts occur in depression (Yamaguchi & Kim, 2013); eating disorders (Goodwin et al., 2014); social anxiety (Kopala-Sibley et al., 2013) and psychosis (White, 2013). As such, self-critical thoughts have been identified as an important treatment target in Cognitive-Behavioral Therapy (CBT) (Fennell, 2006; McManus, Waite & Shafran, 2009) and in more recent 'third-wave therapies' that build on the CBT tradition such as Mindfulness-based interventions (MBI) (e.g. Mindfulness-based Stress Reduction; Mindfulness Based Cognitive Therapy), Acceptance and Commitment Therapy (ACT), and Compassion Focused Therapy (CFT) (Hayes & Hoffman, 2017). These third-wave therapies aim to help people to pay attention 'on purpose, in the present moment, and non-judgmentally' (Kabat-Zinn 1994, p. 4). Although research suggests that third-wave therapies are an effective treatment for multiple psychological disorders (e.g., recurrent depressive disorder; Piet & Hougaard, 2011), consensus has not been reached on what the key change processes in these therapies are (Van der Gucht et al., 2017). As third-wave treatments have become more popular there have been increasing efforts to move beyond standard CBT measures that identify the presence of unhelpful or maladaptive cognitions and supplement them with measures that quantify the way that people relate and respond to internal experiences, including their thoughts.

Gilbert (2009a) suggested that the adoption of harsh and self-critical thoughts in relation to oneself serves as a maladaptive way of defending against criticism from others (Gilbert & Irons, 2005). Like many cognitive experiences, problematic self-criticism will be determined by the context of the experience. The capacity for self-critical thoughts may have evolved as a self-regulatory ability that motivated behavior change in the face of failure or repetition of unhelpful behaviors. Hence, criticizing the self as a discrete action to activate motivation for change may be energizing in particular contexts, but repeated critical self-talk will be demoralizing and likely to undermine effective behavior change efforts. The

suggestion here being that self-critical thoughts serve to limit a person engaging in behaviours that might expose him/her to potentially harmful evaluation from others. The cultivation of self-compassion can be seen as a more adaptive alternative to both self-criticism and low self-esteem (Neff, 2003b). It consists of three main components: self-kindness, common humanity and mindfulness (*ibid.*). Assessment instruments that measure self-compassion (e.g. Self-Compassion Scale: SCS, Neff et al., 2003a) and self-critical thoughts (e.g. Forms of Self-Criticizing/Attacking & Self-Reassuring Scale: FSCRS, Gilbert et al., 2004) have been used to evaluate the efficacy of CFT.

But the mere presence of self-critical thoughts is not necessarily indicative of pathological processes. Self-critical thoughts are intrinsic to human experience and are widely reported in non-clinical samples (Baiao et al., 2014). Indeed, as has been suggested, these may serve a regulatory function (Duarte, Matos et al., 2017). Hence, in addition to measuring the presence and level of self-critical thinking, there is a need to understand how responses to self-critical thoughts may impede engagement in behaviors consistent with personally held values.

Acceptance and Commitment Therapy (ACT; Hayes et al., 1999; 2012) specifically aims to cultivate what has been operationalized as *psychological flexibility* (PF) - "the ability to be in the present moment with full awareness and openness to experiences and to take guided action towards personally held values" (Harris, 2009, p 12). Instead of utilising 'first order' strategies aimed at directly altering the content of thoughts, ACT seeks to explore the functional context in which these experiences occur and employ 'second-order' strategies such as mindfulness, acceptance, or cognitive defusion to enhance PF. Low PF is characterized by behavioral rigidity that stems from efforts to control and suppress difficult internal experiences (e.g. thoughts, feelings, sensations) and is implicated in the development and maintenance of a broad range of psychological problems (Bond & Bunce, 2003) including social anxiety (Dalrymple & Herbet, 2007); depression (Cash & Whittingham, 2010), psychosis (White et al., 2013; 2015) and borderline personality disorder (Rusch et al., 2008). Moreover, studies have also demonstrated that greater PF is positively associated with subjective wellbeing

(A-Tjak et al., 2015; McCracken, Gutierrez-Martinez, & Smyth, 2013; Bohlmeijer et al., 2017).

The most commonly used measure of PF is the *Acceptance and Action Questionnaire* (AAQ-II; Bond et al., 2011), which assesses the extent to which an individual's cognitions can prevent them from engaging in values-consistent actions (e.g. "I'm afraid of my feelings", "My painful memories prevent me from having a fulfilling life"). Because the AAQ-II is very general in its focus a range of context-specific measures of PF have been developed e.g. the Acceptance and Action Questionnaire – Substance Abuse (AAQ-SA; Luoma et al., 2011); Voices Acceptance and Action Questionnaire (V-AAQ; Shawyer et al., 2007); Acceptance and Action Questionnaire for Social Anxiety (AAQ-SA; MacKenzie et al., 2010); the Work-related Acceptance and Action Questionnaire (WAAQ; Bond et al., 2013); Acceptance and Action Questionnaire – Acquired Brain Injury (AAQ-ABI; Whiting et al., 2015); and the Acceptance and Action Questionnaire – Stigma (AAQ-S; Levin et al., 2014).

To date, no measure of PF that focuses specifically on self-critical thoughts has been developed. Given the associations between self-critical thoughts and various forms of mental health difficulties (e.g. depression, social anxiety, eating disorders and psychosis), it is likely that the development of such a measure would have both clinical and research utility. This new measure would complement existing measures that assess the intensity/frequency of self-critical thoughts, but do not measure the extent to which the person becomes psychologically inflexible in response to these thoughts. A measure of this type would have widespread application for third-wave interventions such as MBI, ACT and CFT.

Aims and Hypotheses

The current paper reports on the development of the *Forms of Responding to Self-critical Thoughts Scale* (FoReST), a novel measure of how psychological flexible people are in responding to self-critical thoughts. This measure aims to assess

willingness to experience self-critical thoughts whilst simultaneously committing to values-directed action.

Study 1 entailed the generation of items for the FoReST and an exploratory factor analysis (EFA) of the measure in a convenience non-clinical sample of adults. Study 2 continued the development of the FoReST by conducting a Confirmatory Factor Analysis (CFA). Study 3 explored the construct validity of the FoReST by measuring convergent validity against measures of similar constructs (psychological inflexibility, self-compassion and self-criticism); concurrent validity of the FoReST in relation to theoretically relevant outcomes measures (depression, anxiety and distress), and incremental validity by examining the FoReST's ability to predict levels of depression and anxiety (two clinically important outcomes) beyond an established measure of self-critical thinking (inadequate-self-critical thoughts).

Study 1 Material and Methods

Item generation

The FoReST was developed to assess how psychologically flexible people are in responding to self-critical thoughts rather than the frequency or severity of the self-critical thoughts. It was decided that a single stem statement would be used for each assessment item: 'When I have a critical thought about myself...', with the items themselves taking the form of responses to this stem statement (e.g. "...I try to ignore it").

An initial set of 46 items was generated by the research team by drawing on their own clinical experience and items used in other assessments of psychological inflexibility/flexibility e.g. AAQ-II (Bond et al, 2011) and WAAQ, (Bond et al, 2013). To maximize content validity, the generation of items drew on operational definitions of psychological flexibility/inflexibility that highlighted the importance of both 'acceptance' (i.e. how willing or not a person is to have internal

experiences such as thoughts and emotions) and 'action' (the impact on value congruence and behavioral responding) (Dahl, 2009). Whereas some items were worded to capture a psychologically flexible stance, others were worded to describe a psychologically inflexible stance – thereby ensuring that both poles of a spectrum of psychological flexibility were represented.

A structured focus group consisting of four UK-based Trainee Clinical Psychologists was constituted to evaluate the acceptability, intelligibility and comprehensiveness of the items. At this stage, leading experts in ACT and CFT-related research (Prof. Paul Gilbert and Prof. Dennis Tirch) were also consulted about the content and format of the items. New items or other suggestions about the questionnaire were also invited. On the basis of their feedback, several amendments were made; some items were simplified or abbreviated and three additional items were generated. This version of the FoReST was then tested with three additional UK-based Trainee Clinical Psychologists. Based on their feedback, three poorly worded items were removed.

Procedure

The procedures for study 1 were granted ethical approval by the University of Glasgow, College of Medicine, Veterinary and Life Science, Research Ethics Committee (Ref: 200130039).

A cross-sectional design was used. A Participant Information Sheet (PIS) and assessment scale items were uploaded to SurveyMonkey, a web-based platform designed to host online research data collection and storage. A website developed to promote the research study listed the aims of the study and invited potential participants to click on a link to access the online data collection platform. Contact details for a member of the research team (PL) were provided if participants had any questions.

The PIS indicated that people also had the option of completing the study by meeting in person with a researcher or completing the assessment scales over the

telephone - all participants chose to complete the study online. The assessment battery took approximately 25 minutes to complete. All participants provided informed consent for their involvement.

Recruitment

A number of recruitment strategies were utilized. Advertisements relating to the research were disseminated online via the Facebook pages of relevant UK mental health charities (e.g. Action for Happiness, Mind, Rethink Mental Health), posters were placed in University of Glasgow Student Unions, and an email was circulated to undergraduate students enrolled at the College of Social Sciences, University of Glasgow. All participants were entered into a random draw for a prize valued at £50.

Justification of Sample Size

A sample of 206 participants was recruited by Bond et al. (2011) in their exploratory factor analysis of 49 items in the development of the AAQ-II. The current study followed the 5:1 participant-to-item ratio that has been advocated in previous factor-analytic studies (Gorsuch, 1983; Hatcher, 1994). As there were 46 items in the pool, the research team aimed to recruit a minimum of 250 participants.

Participants

A total of 254 people were recruited to the study. Demographic details for these individuals are provided in Table 1. The vast majority of the participants were female and identified their ethnicity as 'White'. Around one-third of the sample self-reported as a having contact with mental health services either currently (7%) or historically (26%).

INSERT TABLE 1 HERE

Theory/Calculation

Analysis

We conducted an EFA in order to identify one or more latent factors underlying the observed data. We conducted a common factor analysis and used parallel analysis (Horn, 1965) to determine the number of factors to extract. All analyses were conducted on the SPSS statistics program version 22 (IBM Corp, 2013). In accordance with Kaiser's (1960) recommendation, factors with Eigenvalues over 1 were included in the initial model. Existing measures have found unifactorial models of PF (Bond et al., 2012, 2013). However, the possibility could not be ruled out that there could be a multifactorial structure for a measure of how psychologically flexibly people are in responding to self-critical thoughts such as the FoReST. Constituent factors of a superordinate latent construct of a measure of this kind would be expected to correlate so an oblique rotation procedure (Direct Oblimin) was preferred to an orthogonal rotation (Field, 2005). Internal consistency was assessed using SPSS's Reliability function to generate Cronbach's alpha values. Cohen (1992) guidelines were used to interpret effect sizes (ES).

Missing data

Consistent with the approach used by Gillanders et el. (2014) in the development of the cognitive fusion questionnaire (CFQ), where participants missed fewer than three of the FoReST questions, missing scores were prorated using the mean of responses to the remaining FoReST items (N = 18). The remaining missing data were dealt with in the EFA by list-wise deletion, resulting in a final sample of 253 for factor analysis. Several participants completed the FoReST items, but did not complete the additional questionnaires. Participants who responded to less than 90% of items on any of the additional questionnaires were excluded from all validation analyses. The remaining missing responses to each questionnaire were prorated based on the mean of participant's completed responses allowing a sample of 233 for validation analyses.

Results

We first examined the appropriateness of our dataset for analysis. The Kaiser-Meyer-Olkin measure of sampling adequacy and Barlett's test of sphericity both indicated the suitability of the dataset for structure detection. An initial review of the 46 items indicated that seven had low inter-item correlations and eight did not have distributions that approximated normal. After these items were excluded, 31 items remained. At this stage, multicollinearity was found to be an issue as the determinant of the correlation matrix was too low (below 0.00001; Field, 2012). The variance inflation factor (VIF) was assessed to identify problematic items and eleven items consistently had VIF scores above 3, indicating potential multicollinearity. These problematic items were excluded leaving 20 items and an acceptable determinant value (i.e., 0.0000767).

We carried out a series of principal axis factor analyses (FA) on the remaining 20 items. After the initial FA suggested retaining two factors, further FAs were run with a forced 2-factor solution and a Direct Oblimin rotation. Through inspection of the pattern and structure matrixes, we removed items that loaded above .4 on both factors, or below .4 on both factors. After this, we eliminated lower performing items in order to achieve an acceptable level of total variance explained (50%; Streiner, 1994) and to avoid item redundancy (i.e., very high internal consistency). This resulted in a 2-factor 9-item solution (Factor 1 = items 2, 1, 3, 35, 19 and 40; Factor 2 = items 27, 13 and 29) (see Table 2). The Factor 2 items were reverse scored such that higher scores on all the items across both factors indicated more inflexibility in responding to self-critical thoughts. Cronbach's α was at an acceptable level for the total score (Cronbach's α = 0.85), Factor 1 (*unworkable action,* Cronbach's α = 0.86), and Factor 2 (*mindful acceptance*, Cronbach's α = 0.82). A large proportion of variance was accounted for (56%) and there was no issue with multicollinearity (determinant = 0.012). The factors strongly correlated with one another (r = 0.61), but this was not so strong to suggest that they were measuring precisely the same thing. Examination of the item content of the two factors indicated that Factor 1 items coalesced

around the theme of *unworkable action* and Factor 2 items related to *mindful acceptance*.

INSERT TABLE 2 HERE

Study 2 Material and Methods

Procedure

Study Two also used a cross-sectional design. Data collection for this study formed part of an undergraduate student project within the Department of Psychology at Goldsmiths, University of London, UK. Ethical approval was sought and granted from the Department's ethics committee. Participants were undergraduate students studying a range of disciplines from multiple colleges within the University of London. Participants from within Goldsmiths were offered course credit for their participation, whilst students from other colleges at University of London were offered no incentive. Participants were asked to fill in the FoReST trial version as part of a larger pack of measures. They were approached about the study in person or were recruited via project notice boards around the Department of Psychology and once they had expressed interest were sent a link to an electronic version of the questionnaire via the online platform Qualtrics. Informed consent was requested on the first page of the questionnaire and the final page offered debriefing information as well as contact details of the student researcher and their supervisor.

Three participants responded to less than 90% of the FoReST items and were therefore excluded from the analyses. After these exclusions, there were no remaining missing data points.

Participants

A total of 110 participants were recruited. Demographic details for these individuals are provided in Table 3. People who identified as 'White (British/European/Other)' constituted 64% of the sample. The gender of the sample was evenly split between males and females.

INSERT TABLE 3 HERE

Theory/Calculation

Analysis

In our second study, we carried out a CFA to test the fit of the two-factor FoReST model and examine the fit of this model in comparison to an alternative one-factor model where all items were allowed to load onto a single 'Forms of Responding to Self-Critical Thoughts' factor. We used Mplus version 8.2 (Muthén & Muthén, 1998-2017) for all data analysis. We fitted our measurement model using the Maximum likelihood [ML] estimator and evaluated goodness of fit using a combination of absolute and incremental fit indices recommended by Hu and Bentler (1998); specifically, the root mean square error of approximation (RMSEA), the comparative fit index (CFI), the Tucker-Lewis Index (TLI) and standardized root mean square residual (SRMR). We observed the cut-off criteria suggested by Hu and Bentler (1998) which indicate goodness of fit when RMSEA ≤ 0.06, CFI $\geq .95$, TLI ≥ 0.90 and SRMR ≤ 0.08.

Results

Prior to carrying out the CFAs, we ensured the appropriateness of our data for analysis through tests of univariate and multivariate normality. CFA results indicated that despite a slightly poorer fit for the RMSEA, our two-factor FoReST model offered a good overall fit to the data and outperformed the alternative one-

factor model (See Table 4). This indicates that a two-factor model represents the observed data well. All of the unstandardized factor loadings were found to be statistically significant and ranged from 0.70 to 0.96 on Factor 1 (*unworkable action*) and 0.58 to 1.30 on Factor 2 (*mindful acceptance*) (see Table 5). Table 5 also displays the scale means, standard deviations, and internal consistency (which was deemed excellent for Factor 1 and questionable for Factor 2 according to George and Mallery's (2003) criteria).

INSERT TABLE 4 HERE

INSERT TABLE 5 HERE

Study 3 Material and Methods

Procedure

Study Three was conducted to explore the convergent, concurrent and incremental validity of the FoReST. Data from Study One (i.e. Sample 1) and Study Two (i.e. Sample 2) were used to assess the convergent validity of the FoReST by assessing correlations with the AAQ-II (Sample 1 and 2), the Self-Compassion Scale (SCS) (Sample 1 and 2) and the Forms of Self-Criticizing/Attacking & Self-Reassuring Scale (FSCRS) subscales (Sample 1 only). The concurrent validity of the FoReST was assessed by investigating correlations with the subscales of the Hospital Anxiety and Depression Scale (HADS) (Sample 1) and the General Health Questionnaire (GHQ-12) (Sample 2). In addition, we used Sample 1 to examine whether there were significant differences on the FoReST between participants scoring above the cut-off for severe levels of depression on the HADS-Depression and HADS-Anxiety (scores ≥ 8) compared to those scoring below the cut-off (score < 8. Finally, incremental validity was assessed by determining if the FoReST accounted for significantly greater proportion of the variance in HADS-Depression and HADS-Anxiety subscale scores (Sample 1) compared to other variables including age/gender and levels of self-critical thoughts (using the 'inadequate-self' subscale of the FSCRS).

Measures

Acceptance and Action Questionnaire (AAQ-II; Bond et al., 2011): The AAQ-II is a 7-item scale that measures psychological inflexibility. When originally validated it demonstrated good internal consistency (α = 0.84), test-retest reliability (r = 0.79), and construct validity (Bond et al., 2011). In the current study, the internal consistency of the AAQ-II in Sample 1 was α = 0.93 and in Sample 2 it was α = 0.84

Forms of Self-Criticizing/Attacking & Self-Reassuring Scale (FSCRS; Gilbert et al, 2004): The FSCRS is a 22-item scale assessing the level and form of participants' self-critical and self-reassuring thoughts (e.g. "when things go wrong for me I am easily disappointed with myself"). The 'Inadequate-Self' and 'Self-Hating' subscales were found to have very good internal consistency of (α = .90 and .86 respectively) in a sample of female students (Gilbert et al, 2004). The internal consistency of the 'Inadequate-Self', 'Self-Hating' and 'Reassuring-Self' subscales of the FSCRS in Sample 1 in the current study were α = 0.92, α = 0.84 and α = 0.93 respectively.

Self-Compassion Scale (SCS; Neff et al., 2003): The SCS is a 26-item measure that measures trait level self-compassion. It includes items that measure how frequently people respond to feelings of inadequacy or suffering with self-kindness, self-judgment, common humanity, isolation, mindfulness, and over-identification. It has been shown to have excellent internal consistency in a student sample (α= 0.92) (Neff et al, 2003). In the current study, the internal consistency of the SCS in Study 1 was also excellent (α= 0.96). The Self-Compassion Scale (SCS-SF; Raes et al., 2011), a 12-item version of the scale, was used in Study 2 (α= 0.96)

Hospital Anxiety and Depression Scale (HADS; Snaith and Zigmond, 1994): The HADS is a 14-item measure of levels of anxiety and depression symptomatology (7-items each). HADS-A has demonstrated levels of internal consistency between α = 0.68 to 0.93 (mean α = 0.83) and for HADS-D scored between α = 0.67 to 0.90

(mean α = 0.82) (Bjelland et al., 2001). In Sample 1 of the current study, the internal consistency of the HADS-A was α= .87 and the HADS-D was α= .85.

General Health Questionnaire (GHQ-12; Goldberg & Williams, 1988) is a 12-item measure that was developed as a screen for non-specific mental health difficulties. Research evidence has indicated that the measure is valid (Harding et al., 1999). In the current study the 'Likert 'scoring system was used to allocate responses a score of either 0, 1, 2 or 3. The internal consistency of the GHQ-12 in Study 2 was shown to be excellent (α= 0.92). Over half of the participants (57%) in Study 2 met criteria for 'caseness' for common mental health difficulties (i.e. GHQ-12 score ≥ 12).

Analysis

Boxplots indicated possible outliers for several of the additional measures indicating that parametric correlations may not be appropriate. For this reason, the more conservative *Spearman's* ρ analyses were conducted. A *t-test* was conducted to determine if there were significant differences on the FoReST between Sample 1 participants scoring above the cut-off for severe levels of depression on the HADS-Depression (score ≥ 8) compared to those scoring below the cut-off (score < 8). Consistent with Hunsley and Meyer's (2003) stringent approach to incremental validity, hierarchical linear regression analyses were undertaken with Sample 1 to determine if the FoReST made a statistically significant contribution to the proportion of the variance in HADS-Depression and HADS-Anxiety scores beyond demographic variables (i.e. age/gender) and other relevant constructs (i.e. levels of 'inadequate-self' FSCSR). Depression and anxiety were focused on for these analyses as these conditions are recognized as the most burdensome of mental disorders (Whiteford et al., 2015) and are collectively referred to as 'common mental disorders'. Levels of self-critical thoughts (as assessed by the 'inadequate self' subscale of the FSCSR) have also been shown to be highly associated with levels of depression (Gilbert et al., 2006).

Results

Table 6 provides details of the mean scores and standard deviations for the various assessment instruments. In addition, correlation coefficients of associations between FoReST and other relevant assessment measures are provided. In sample 1, the FoReST was significant correlated with the AAQ-II (ρ = 0.73, $p \leq 0.001$), SCS (ρ = -0.80, $p \leq 0.001$), FSCRS Reassured Self (ρ = -0.72, $p \leq 0.001$), FSCRS Hated Self (ρ = 0.73, $p \leq 0.001$), FSCRS Inadequate Self (ρ = 0.71, $p \leq 0.001$), HADS Depression Subscale (ρ = 0.59, $p \leq 0.001$) and HADS Anxiety Subscale (ρ = 0.57, $p \leq 0.001$). In Sample 2, the FoReST showed significant correlations with the AAQ-II (ρ = 0.84, $p \leq 0.001$), SCS-SF (ρ = -0.87, $p \leq 0.001$), and General Health Questionnaire-12 (ρ = 0.70, $p \leq 0.001$). Each of the two FoReST factors demonstrated similar patterns of association with the other measures in the two samples.

INSERT TABLE 6 HERE

A *t-test* analysis indicated that the FoReST mean scores (X = 42.00, S.D. = 5.10) of those meeting caseness (score ≥8) on the HADS-Depression subscale (n = 34) were significantly higher than the mean scores (X=30.03, S.D. = 8.33) those of individuals that did not meet caseness (n = 208) (t = -11.415, df = 65.944, $p < 0.001$). Similarly, the FoReST mean scores (X = 36.48, S.D. = 7.67) of those meeting caseness (score ≥8) on the HADS-Anxiety subscale (n = 113) were significantly higher than the mean scores (X=27.22, S.D. = 7.81) those of individuals that did not meet caseness (n = 130) (t = -9.28, df = 241, $p < 0.001$).

To measure incremental validity, hierarchical regression analyses was conducted with Sample 1 data to ascertain the effect on levels of depression (HADS-Depression) and anxiety (HADS-Anxiety) of: gender and age, the level of "inadequate-self" critical thoughts, and how psychologically flexible the person is in responding to self-critical thoughts.

With regard to levels of depression (see: Table 7), there was no violation of independence of errors (Durban-Watson = 1.968) and no concerns about multicollinearity (VIF Range = 1.005 to 1.902). The first step of the regression consisted of gender and age, levels of "inadequate-self" critical thoughts was added at the second step, with how psychologically flexible the person is in responding to self-critical thoughts added at the final stage. The overall regression model predicted approximately 43% of variance in depression (R^2 = 0.43, F(4,229) = 43.074, p<0.001). After controlling for age and gender, step 2 which saw the addition of levels of 'inadequate-self' critical thoughts, predicted 32% of variance of depression scores. Finally controlling for age, gender, and levels of inadequate-self thoughts, step 3 which focused on the contribution of how psychologically flexible people are in responding to self-critical thoughts predicted approximately 10% of variance in depression scores. The change in R^2 between steps 2 and 3 was statistically significant.

With regard to levels of anxiety (see: Table 8), there was no violation of independence of errors (Durban-Watson = 2.133) and no concerns about multicollinearity (VIF Range = 1.004 to 1.881). As with the previously reported regression analysis, the first step of the regression consisted of gender and age, levels of "inadequate-self" critical thoughts was added at the second step, with how psychologically flexible the person is in responding to self-critical thoughts added at the final stage. The overall regression model predicted approximately 47% of variance in anxiety (R^2 = 0.47, F(4,230) = 51.53, p<0.001). After controlling for age and gender, step 2 which saw the addition of levels of 'inadequate-self' critical thoughts, predicted 39% of variance of anxiety scores. Finally controlling for age, gender, and levels of inadequate-self thoughts, step 3 which focused on the contribution of how psychologically flexible people are in responding to self-critical thoughts predicted approximately 3% of variance in anxiety scores. The change in R^2 between steps 2 and 3 was statistically significant.

Discussion

In this series of studies, we systematically developed and examined the psychometric properties of a novel measure of how psychologically flexible people are in responding to self-critical thoughts (the FoReST). A measure of this type should have applications with third-wave psychological interventions (e.g. ACT, MBI and CFT) for mental health difficulties where self-critical thoughts are a prominent causal or maintenance factor (e.g. depression, eating disorders, psychosis, social anxiety etc). Study One derived an acceptable 2-factor (9-item) model for the FoReST. This model explained approximately 56% of variance. According to George and Mallery's (2003) criteria, the measure demonstrated good internal consistency overall, good internal consistency for factor 1 and good internal consistency for factor 2.

Factor analytic studies of other measures of PF have set out with similar operational definitions to the FoReST and found unifactorial solutions to be the strongest (e.g. AAQ-II, Bond et al, 2011; WAAQ, Bond et al, 2013). However, in terms of content validity (i.e. the extent to which a measure represents all facets of a given construct), the 9-items of the FoReST which include a focus on both *unworkable action* and *mindful acceptance* fits with previous definitions of PF (see Dahl, 2009).

The confirmatory factor analysis of the 9-item FoReST in a sample of undergraduate university students in Study Two shows that the two-factor solution of '*unworkable action*' (Factor 1) and '*mindful acceptance*'' (Factor 2) showed good fit to the data. In this sample, the FoReST showed an excellent level of overall internal consistency (Cronbach's α = 0.91). Although the internal consistency for Factor 1 was excellent in Study 2, the internal consistency for Factor 2 was questionable according to George and Mallery's (2003) criteria. However, this questionable level of internal consistency for this factor in this particular sample is not an unusual occurrence in the early stages of test development (Nunnally, 1978).

The results of Study Three provided support for the convergent and predictive validity of the FoReST across two separate samples. In terms of convergent validity, the FoReST had significant correlations in the expected directions with the measures of psychological inflexibility, self-compassion and self-critical thoughts. The correlational co-efficients for the associations that the FoReST had with the AAQ-II and the SCS were particularly high. This could suggest that there is considerable conceptual overlap between the concepts. In the current study the SCS was correlated negatively and significantly with the AAQ-II ($r = -0.735$, $p < 0.001$) in Study 1, and the SCS-SF was correlated negatively and significantly with the AAQ-II ($r = -0.846$, $p < 0.001$) in Study 2. These findings are consistent with previous research conducted by Marshall and Brockman (2013) that focused on 144 university students in Australia. They found that the SCS-SF was significantly negatively correlated with the AAQ-II ($r = -0.574$, $p < 0.001$). It has been suggested that the strong associations between self-compassion and psychological flexibility are due to a person needing to 'hold a painful internal experience' before a kind and soothing response can be initiated (Marshall and Brockman, 2013).

With regard to concurrent validity, there were significant correlations in the expected directions with measures relating to mental health (including depression, anxiety and distress). Future research would benefit from investigating the association between the FoReST and relates to other outcome measures assessing constructs such as job, life and/or relationship satisfaction, subjective wellbeing and quality of life.

When participants were categorized according to their level of HADS-Depression and HADS-Anxiety scores, the difference in the FoReST scores between the groups was statistically significant and in the expected direction. Hierarchical regression analyses demonstrated the predictive validity of forms of responding to self-critical thoughts (assessed by the FoReST) for predicting variance in depression and anxiety beyond the contribution of demographic characteristics and an established predictor of depression and anxiety (i.e. self-critical thoughts). This suggests that people's inability to be psychologically flexible in responding to self-

critical thoughts when these thoughts occur, rather than the presence of these thoughts alone, is associated with increased levels of depression and anxiety.

In addition to ACT and MBI, the findings of the current study have important implications for *Compassion Focused Therapy* (CFT). CFT seeks to help individuals develop a compassionate stance towards themselves and to cope with challenging emotions with a greater degree of understanding, self-directed care, and openness (Gilbert, 2009b). As yet, however, there are no specific measures assessing the extent to which people are psychologically flexible in responding to self-critical thoughts. A key driver in the development of the FoReST was the extent to which MBI, ACT and CFT potentially overlap in terms of processes of change, and the potential for developing integrative approaches that combine aspects of these third-wave approaches. The FoReST potentially represents an important process of change measure for these efforts.

Limitations

There were a number of potential limitations associated with Study One. The items of the FoReST are not intended to individually assess each of the six processes that the Hexaflex model of ACT (Hayes et al., 2012) proposed to contribute to psychological flexibility. Instead, the intention was to develop items that reflect the overarching experience of psychological flexibility. It is possible that attending to specific aspects of the ACT model may have served to generate a broader range of potential items for assessing flexible responding to self-critical thoughts. However, it is noteworthy that more recent iterations of the ACT model such as the *Triflex* approach (Harris, 2009) propose that *Being Present*, *Being Open* and *Doing What Matters* represent constituent aspects of psychological flexibility. We are confident that the items of the FoReST have good face validity for broadly assessing psychological flexibility in relation to self-critical thoughts. Although members of the research team drew on their clinical experience of working with people presenting with problematic self-critical thinking when generating the initial item-pool for the FoReST, service-users were not formally consulted during

this process. People who have difficulties rooted in self-critical thoughts would have given unique insights to develop the content and acceptability of the FoReST.

Efforts were made to recruit a range of different participants, but the sample was one of convenience rather than a representative sample. The vast majority of the participants in Study One were female and of white ethnicity. In Study Two the representation of different genders was more balanced, and there was more representation from minority groups. Research has indicated that many widely used measures perform differently across ethnic groups (Hambrick et al, 2010), therefore collecting further data from participants of diverse ethnicities would have been advantageous. As with Study One, the samples recruited to Study Two (and hence Study Three) were recruited through convenience sampling – this time focusing exclusively on a student population. The FoReST is most likely to be employed as an instrument to assess the efficacy of third-wave interventions in clinical samples, so the absence of participants being specifically recruited from clinical services in the studies reported in the current paper could be viewed as a limitation. It is noteworthy that over half of participants (57%) in Study 2 met caseness criteria for common mental health problems on the GHQ-12, and we contend that testing the FoReST in heterogenous samples (encompassing varying levels of the target construct) is an important initial step. This is also consistent with the adiagnostic approach that practitioners of therapies such as Acceptance and Commitment Therapy (and other third wave therapies) often advocate for.

A final potential limitation relates to the wording of some of the FoReST items that rely on phrasing and terminology which are colloquial to those familiar with third-wave interventions and may be less clear to those who are not acquainted with these approaches (e.g. “get caught up in it” or “let it pass from my awareness”). This may have contributed to the questionable internal consistency of the ‘mindful acceptance’ items in Sample 2. This is worthy of further investigation in future studies.

Future Research

Although the sample sizes recruited in the current studies were sufficiently large to support the CFA and EFA, future research should seek to replicate the factor structure and investigate the psychometric properties of the FoReST in normative and also clinical samples. This should also provide important opportunities to investigate further the internal consistency of Factor 2 of the FoReST, which although good in Sample 1, was adjudged to be questionable in Sample 2. One potential explanation for the questionable internal consistency of Factor 2 in Sample 2 could be the low number of items it is comprised of (i.e. three items). Indeed, it has been found that Cronbach alpha estimates increase as scale length increases (Voss, Stem, & Fotopoulos, 2000).

The fact that both Study One and Study Two used a cross-sectional design meant that it was not possible to explore how changes in psychological flexibility about self-critical thoughts was associated with changes in standardized measures assessing similar constructs and psychological distress over time. Conducting longitudinal studies is therefore recommended. This would help determine the test-retest reliability of the FoReST as well as assessing the consistency of the FoReST scores over longer periods of time. In terms of additional specific settings in which the FoReST could be utilised, workplace psychology research has been interested in the role of concepts that aspects of 'self-relating' (e.g., self-efficacy, self-esteem, self-concept clarity etc.) have on work performance (Bowling, Eschleman, Wang, Kirkendall, & Alarcon, 2010; McIntyre, Mattingly, Lewandowski & Simpson, 2014). Future research could also explore the extent to which psychological flexibility relating to self-critical thoughts may impact on aspects of work performance and/or levels of burnout.

Conclusion

Our data provides preliminary support for the coherence, validity, and internal consistency of the FoReST – an easy to administer measure that assesses people's tendency to act in a closed, inflexible, values-incongruent manner, in the presence

of self-critical thoughts. The FoReST has potential clinical applications with people experiencing problematic levels of shame and guilt; which can include, but is not limited to, those with a lived experience of depression, eating disorders, social anxiety and psychosis. The FoReST provides important opportunities for tracking changes in a potentially relevant process of change in the delivery of psychological interventions – in particular third-wave interventions such as MBI, ACT, and CFT. Further research is required to assess the psychometric properties of the FoReST in clinical and normative samples.

References

A-Tjak, J.G., Davis, M.L., Morina, N., Powers, M.B., Smits, J.A., & Emmelkamp, P.M. (2015). A meta-analysis of the efficacy of acceptance and commitment therapy for clinically relevant mental and physical health problems. Psychotherapy and psychosomatics, 84 (1), 30–6.

Bjelland, I., Dahl, A.A., Haug, T.T., Neckelmann, D., 2002. The validity of the Hospital Anxiety and Depression Scale: an updated literature review. J. Psychosom. Res. 52, 69–77.

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

Bond, F. W. & Bunce, D. (2003). The role of acceptance and job control in mental health, job satisfaction and work performance. Journal of Applied Psychology, 88, 1057-1067.

Bond, F. W., Hayes, S. C., Baer, R. A., Carpenter, K. M., Guenole, N., Orcutt, H. K., & Zettle, R. D. (2011). Preliminary psychometric properties of the Acceptance and Action Questionnaire–II: A revised measure of psychological inflexibility and experiential avoidance. Behavior Therapy, 42(4), 676-688.

Bond, F. W., Lloyd, J., & Guenole, N. (2013). Preliminary psychometric properties of the work-related acceptance and action questionnaire. Journal of Occupational and Organizational Psychology, 86, 331–347.

Bowling, N. A., Eschleman, K. J., Wang, Q., Kirkendall, C., & Alarcon, G. (2010). A meta‐analysis of the predictors and consequences of organization‐based self‐esteem. Journal of Occupational and Organizational Psychology, 83(3), 601-626.

Cash, M., & Whittingham, K. (2010). What facets of mindfulness contribute to psychological well-being and depressive, anxious, and stress related symptomology? Mindfulness, 1, 177-182.

Ciarrochi, J., Kashdan, T. B., & Harris, R. (2013). The Foundations of Flourishing. In T. B, Kashdan & J. Ciarrochi (Eds.), Mindfulness, acceptance and positive psychology (pp 1-29). Oakland: Context Press.

Costello, A.B. & Osborne, J.W. (2005). Best Practices in Exploratory Factor Analysis: Four Recommendations for Getting the Most From Your Analysis. *Practical Assessment Research & Evaluation*, 10, 1-9.

Credé, M., Chernyshenko, O. S., Bagraim, J., & Sully, M. (2009). Contextual performance and the job satisfaction–dissatisfaction distinction: Examining artifacts and utility. Human performance, 22(3), 246-272.

Dahl, J., Lundgren, T., Plumb, J., & Stewart, I. (2009). The Art and Science of Valuing in Psychotherapy: Helping Clients Discover, Explore, and Commit to Valued Action Using Acceptance and Commitment Thera. New Harbinger Publications.

Dalrymple, K. L., & Herbert, J. D. (2007). Acceptance and commitment therapy for generalized social anxiety disorder. A pilot study. Behavior Modification, 31, 543−568.

Duarte, C., Matos, M., Stubbs, R. J., Gale, C., Morris, L., Gouveia, J. P., & Gilbert, P. (2017). The impact of shame, self-criticism and social rank on eating behaviours in

overweight and obese women participating in a weight management programme. PloS one, 12(1), e0167571.

Elahi, A., Algorta, G. P., Varese, F., McIntyre, J. C., & Bentall, R. P. (2017). Do paranoid delusions exist on a continuum with subclinical paranoia? A multi-method taxometric study. Schizophrenia Research.
https://doi.org/10.1016/j.schres.2017.03.022

Fennell, M. (2006). Overcoming low self-esteem: Self-help program. London: Constable and Robinson.

Fledderus, M., Bohlmeijer, E. T., Pieterse, M. E., & Schreurs, K. M. G. (2012). Acceptance and commitment therapy as guided self-help for psychological distress and positive mental health: a randomized controlled trial. Psychological medicine, 42(3), 485-495.

Faul, F., Erdfelder, E., Buchner, A., & Lang, A.-G. (2009). Statistical power analyses using G*Power 3.1: Tests for correlation and regression analyses. *Behavior Research Methods*, *41*, 1149-1160.

Field, A. (2005). Discovering statistics using IBM SPSS statistics. Sage.

Field, A. (2013). Discovering statistics using IBM SPSS statistics. Sage.

Forman, E.M., Herbert, J.D., Juarascio, A.S., Yeomans, P.Y., Zebell, J.A., Goetter, E.M. & Moitra, E. (2012). The Drexel defusion scale: A new measure of experiential distancing. *Journal of Contextual Behavioral Science, 1,* 55-65.

George, D., & Mallery, P. (2003). SPSS for Windows step by step: A simple guide and reference. 11.0 update (4th ed.). Boston: Allyn & Bacon

Goldberg, D., & Williams R. (1988). A user's guide to the General Health Questionnaire. NFER-Nelson, Windsor

Goodwin, H., Arcelus, J., Geach, N., & Meyer, C. (2014). Perfectionism and eating psychopathology among dancers: The role of high standards and self-criticism. *European Eating Disorders Review, 22*(5), 346-351.

Gilbert, P. (2009a). Introducing compassion-focussed therapy. Advances in Psychiatric Treatment, 15, 199 – 208.

Gilbert, P. (2009b). The Compassionate Mind. London: Constable Robinson and USA: New Harbinger.

Gilbert, P., Clarke M, Hempel S, Miles J.N.V.& Irons C. (2004). Criticizing and reassuring oneself: An exploration of forms style and reasons in female students. British Journal of Clinical Psychology 43, 31-50.

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., 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(2), 183.

Gillanders, D.T., Bolderston, H., Bond, F.W., Dempster, M., Flaxman, P.E, Campbell, L., Tansey, L., Noel, P., Ferenback, C., Masley, S., Roach, L., Lloyd, J, Clarke, S., & Remmington, B. (2014). The Development and Initial Validation of the Cognitive Fusion Questionnaire. Behavior Therapy, 45, 83-101.

Gorsuch R. L. (1983) Factor Analysis (2nd edn.). Erlbaum. Hillsdale. N.J.

Hambrick, J.P., Rodebaugh, T.L., Balsis, S., Woods, C.M., Mendez, J.L., & Heimberg, R.G. (2010). Cross-ethnic measurement equivalence of measures of depression, social anxiety, and worry, Assessment, 17,155–171.

Hardy GE, Shapiro DA, Haynes CE, Rick JE. (1999). Validation of the General Health Questionnaire-12: Using a sample of employees from England's health care services. Psychological Assessment. 11, 159–165.

Harris, R. (2009). ACT Made Simple: An Easy-To-Read Primer on Acceptance and Commitment Therapy. Oakland, CA: New Harbinger.

Hatcher, L. (1994). A step-by-step approach to using the SAS system for factor analysis and structural equation modeling, Cary, NC: SAS Institute.

Hayes, S. C., Strosahl, K. D., & Wilson, K.G. (1999). Acceptance and commitment therapy: An experiential approach to behavior change. New York, NY: The Guilford Press.

Hayes, S. C., Strosahl, K. D., & Wilson, K. G. (2012). Acceptance and commitment therapy: The process and practice of mindful change (2nd edition). New York, NY: Guilford Press.

Hayes, S. C., & Hofmann, S. G. (2017). The third wave of cognitive behavioral therapy and the rise of process‐based care. *World Psychiatry*, *16*(3), 245-246.

Hunsley, J., & Mash, E. J. (2007). Evidence-based assessment. *Annual Review of Clinical Psychology*, 3, 29–51.

IBM Corp. (2013). IBM SPSS Statistics for Windows, Version 22.0. Armonk, NY: IBM Corp.

Kaiser, H. F. (1960). The application of electronic computers to factor analysis. Educational and Psychological Measurement, 20, 141-151.

Kopala-Sibley, D.C., Zuroff, D.C., Russell, J.J., & Moskowitz, D.S. (2013). Understanding heterogeneity in social anxiety disorder: Dependency and self-criticism

moderate fear responses to interpersonal cues. British Journal of Clinical Psychology, 53(2), 141-156.

Levin, M. E., MacLane, C., Daflos, S., Seeley, J. R., Hayes, S. C., Biglan, A. & Pistorello, J. (2014). Examining Psychological Inflexibility as a Transdiagnostic process across psychological disorders. Journal of Contextual Behavioral Science, 3, 155–163.

Luoma, J., Drake, C.E., Kohlenberg, B.S., & Hayes, S. C. (2011). Substance abuse and psychological flexibility: The development of a new measure. Addiction Research and Theory, 19(1), 3-13.

MacKenzie, M. B., & Kocovski, N. L. (2010). Self-reported acceptance of social anxiety symptoms: Development and validation of the Social Anxiety—Acceptance and Action Questionnaire. International Journal of Behavioral Consultation and Therapy, 6(3), 214.

Marshall, E.-J. & Brockman, R.N. (2016). The Relationships Between Psychological Flexibility, Self-Compassion, and Emotional Well-Being. Journal of Cognitive Psychotherapy: An International Quarterly, 30, 60-72.

McCracken, L. M., Gutiérrez-Martínez, O., & Smyth, C. (2013). "Decentering" Reflects Psychological Flexibility in People With Chronic Pain and Correlates With Their Quality of Functioning. Health Psychology, 32(7), 820-823.

McIntyre, K. P., Mattingly, B. A., Lewandowski Jr, G. W., & Simpson, A. (2014). Workplace self-expansion: Implications for job satisfaction, commitment, self-concept clarity, and self-esteem among the employed and unemployed. Basic and Applied Social Psychology, 36(1), 59-69.

McManus, F., Waite, P., & Shafran, R. (2009). Cognitive-behavior therapy for low self-esteem: A case example. Cognitive and Behavioral Practice, 16(3), 266-275.

Mundt, J. C., I. M. Marks, et al. (2002). "The Work and Social Adjustment Scale: a simple measure of impairment in functioning." British Journal of Psychiatry,180, 461-464.

Neff, K. D. (2003a). Self-compassion: An alternative conceptualization of a healthy attitude toward oneself. Self and Identity, 2, 85–101.

Neff, K. D. (2003b). The development and validation of a scale to measure self-compassion. Self and Identity, 2(3), 223-250.

Nunnally, J. C. (1978). An overview of psychological measurement. In *Clinical diagnosis of mental disorders* (pp. 97-146). Springer, Boston, MA.

Nunnally, J. C., Bernstein, I. H. (1994). Psychometric theory (3rd ed.). New York: McGraw Hill.

Piet, J., & Hougaard, E. (2011). The effect of mindfulness-based cognitive therapy for prevention of relapse in *recurrent major depressive disorder: a systematic review and meta-analysis. Clinical Psychology Review*, 31(6), 1032-1040.

Raes, F., Pommier, E., Neff,K. D., & Van Gucht, D. (2011). Construction and factorial validation of a short form of the Self-Compassion Scale. *Clinical Psychology & Psychotherapy*. 18, 250-255.

Rüsch, N., Schiel, S., Corrigan, P. W., Leihener, F., Jacob, G. A., Olschewski, M., et al. (2008). Predictors of dropout from inpatient dialectical behavior therapy among women with borderline personality disorder. *Journal of Behavior Therapy and Experimental Psychiatry*, 39, 497−503.

Sandoz E. K., Wilson K. G., Merwin R. M., & Kellum K. K. (2013). Assessment of body image flexibility: the body image-acceptance and action questionnaire Journal of Contextual and Behavioural Science, 2, 39–48.

Shawyer, F., Ratcliff, K., Mackinnon, A., Farhall, J., Hayes, S. C., & Copolov, D. (2007). The voices acceptance and action scale (VAAS): Pilot data. Journal of clinical psychology, 63(6), 593-606.

Smout, M., Davies, M., Burns, N., & Christies, A. (2014). Development of the Valuing Questionnaire (VQ). Journal of Contextual Behaviral Science, 3, 164-172.

Snaith, R. P., & Zigmond, A. S. (1994). HADS: Hospital Anxiety and Depression Scale. 1994. Windsor: NFER Nelson.

Tucker, L. & MacCallum, R. (1997). Exploratory factor analysis: A book manuscript. Retrieved 30th September 2017 from http://www.unc.edu/~rcm/book/factornew.htm.

Van der Gucht, K., Takano, K., Raes, F., & Kuppens, P. (2017). Processes of change in a school-based mindfulness programme: cognitive reactivity and self-coldness as mediators. *Cognition and Emotion*, 1-8.

Voss, K. E., Stem, D. E., & Fotopoulos, S. (2000). A comment on the relationship between coefficient alpha and scale characteristics. *Marketing Letters*, 11(2), 177-191.

Yamaguchi, A., & Kim, M. (2013). Effects of Self-Criticism and Its Relationship with Depression Across Cultures. *International Journal of Psychological Studies, 5*(1), 1-10.

White, R. G. (2013). Treating depression in psychosis: self-compassion as a valid life direction. In B. Gaudiano (Ed.) Incorporating Acceptance and Mindfulness into the treatment of Psychosis: Current Trends and Future Directions. Oxford University Press.

White, R. G., Gumley, A. I., McTaggart, J., Rattrie, L., McConville, D., Cleare, S., & Mitchell, G. (2012). Depression and anxiety following psychosis: associations with mindfulness and psychological flexibility. Behavioural and Cognitive Psychotherapy, 1, 1–18.

White, R.G., Gumley, A.I., McTaggart, J., Rattrie, L., McConville, D., Cleare, S., & Mitchell, G. (2015). Acceptance and commitment therapy for depression following psychosis: An examination of clinically significant change. Journal of Contextual Behavioral Science, 4, 203–209.

Whiteford, H. A., Ferrari, A. J., Degenhardt, L., Feigin, V., & Vos, T. (2015). The global burden of mental, neurological and substance use disorders: an analysis from the Global Burden of Disease Study 2010. PloS one, 10(2), e0116820.

Whiting, D. L., Deane, F. P., Ciarrochi, J., McLeod, H. J., & Simpson, G. K. (2015). Validating measures of psychological flexibility in a population with acquired brain injury. Psychological Assessment, 27(2), 415–423.

Declarations of interest: none.

Acknowledgements

The authors would like to acknowledge the helpful support that Prof Paul Gilbert, Dr Dennis Tirch, Dr Nic Hooper, Dr Jason Luoma, Dr Paul Christiansen, Dr Alex Stenhoff provided in the development of the FoReST.

Table 1. Demographic characteristics of participants in Study One

		Participants (N= 254)	*(%)*
Age			
	Mean = 31.0 years (SD= 9.0)		
Gender			
	Male	54	21.3
	Female	198	77.9
	Transgender	2	0.8
	Total*	254	100.0
Ethnicity			
	White	237	94.8
	Other	13	5.2
	Total*	250	100.0
Employment			
	Employed	188	74.6
	Student	52	20.6
	Not employed/retired	12	5.2
	Total*	252	100.0
Relationship status			
	Married	90	36.0
	Long-term relationship	86	34.4
	Single/divorced	74	29.6
	Total*	250	100.0
Mental health assessed/treated			
	Currently	18	7.3
	Historically	65	26.2
	Never	165	66.5
	Total*	248	100
Practice mindfulness/meditation			
	Regularly (daily – weekly)	35	13.9
	Monthly or less	65	25.8
	Never	152	60.3
	Total*	252	100

*Some participants did not report all demographic data (N for EFA was 254)

Table 2. Factor Loadings of FoReST items in Study One (N= 253)

Item	2-Factor Solution	
	Factor 1	Factor 2
"When I have a critical thought about myself….."		
2. I act in a way that makes life more difficult for me	**.845**	-.023
1. It gets me so down that I don't act in the way I should	**.744**	.117
3. I don't treat myself the way I would like	**.686**	.186
35. I don't treat others the way I would like	**.685**	-.101
19. I don't try as hard	**.601**	-.007
40. I waste more of my time	**.577**	.068
27. I can let it pass from my awareness in its own time	-.049	**.941**
13. I can let the feelings it creates pass from my awareness in their own time	.000	**.716**
29. I notice it without getting too caught up in it	.161	**.611**
Scale Mean	20.47	10.94
Standard deviation	6.49	3.51
% Variance Explained	47.7%	8.30%
TOTAL % Variance Explained	56%	
Internal consistency (Cronbach α)	.86	.82
Internal consistency overall (Cronbach α)	.85	

N.B. Entries in bold represent the factor on which the item loaded most highly.

Table 3. Demographic characteristics of participants in Study Two

		Participants (N= 110)	*(%)*
Age			
	Mean = 21.55 years (SD = 2.96)		
Gender			
	Male	54	49.0
	Female	55	50.0
	Non-binary	1	1.0
Ethnicity			
	White British	50	44.2
	White European	11	10.0
	White Other	11	10.0
	Mixed Ethnicity	6	5.5
	Asian/Asian British	14	12.7
	Black African/Caribbean	16	14.5
	Middle Eastern	2	1.8
Area of study			
	Psychology Related	66	60.0
	Finance/Business Related	15	13.6
	Social Sciences	14	12.7
	Art	6	5.5
	Drama	4	3.6
	Music	5	4.5

Table 4 Confirmatory factor analysis results for the FoReST from Study Two

	Models	χ^2	*df*	CFI	TLI	RMSEA	SRMR	$\Delta\chi^2$ *against Model 1*	Δdf *against Model 1*
1.	Two-factor	48.81**	26	0.97	0.96	0.09	0.04	-	-
2.	One-factor	78.47***	27	0.93	0.91	0.13	0.06	29.66	1[a]

Note. N = 110; FoReST = The Flexibility of Responses to Self-Critical Thoughts Scale; χ^2 = chi-square value; *df* = degrees of freedom; CFI = comparative fit index; TLI = Tucker-Lewis index; RMSEA = root mean square error of approximation; SRMR = standardized root mean square residual; [a] = the comparison model offered a significantly worse fit than Model 1; ** $p < .01$, *** $p < .001$.

Table 5. Unstandardized factor loadings from confirmatory factor analyses, means, standard deviations and alpha coefficients for the FoReST in Study Two

	FoReST scale loadings	
Item	Factor 1	Factor 2
"When I have a critical thought about myself....."		
2. I act in a way that makes life more difficult for me	1.00	-
1. It gets me so down that I don't act in the way I should	0.90	-
3. I don't treat myself the way I would like	0.96	-
35. I don't treat others the way I would like	0.70	-
19. I don't try as hard	0.73	-
40. I waste more of my time	0.91	-
27. I can let it pass from my awareness in its own time	-	1.00
13. I can let the feelings it creates pass from my awareness in their own time	-	1.30
29. I notice it without getting too caught up in it	-	0.58
Scale mean	29.97	12.15
Standard deviation	8.69	2.81
Internal consistency scales (Cronbach α)	.93	.64
Internal consistency overall (Cronbach α)	.91	

Note. $N = 110$; FoReST = The Flexibility of Responses to Self-Critical Thoughts Scale; All factor loadings are significant at $p < .001$.

Table 6. Correlation coefficients of associations between FoReST and other relevant assessment measures

Measure	Sample	N	Mean (SD)	Correlation (Spearman's ρ)		
				Factor 1 – *Unworkable Action*	Factor 2 – *Mindful Awareness*	FoReST Total
FoReST Factor 1	1	254	20.47 (6.49)	N/A	0.61***	0.95***
FoReST Factor 1	2	110	18.03 (8.69)	N/A	0.59***	0.98***
FoReST Factor 2	1	254	10.95 (3.52)	0.61***	N/A	0.82***
FoReST Factor 2	2	110	11.85 (2.81)	0.59***	N/A	0.75***
FoReST Total	1	254	31.42 (9.02)	0.95***	0.82***	N/A
FoReST Total	2	110	29.88 (10.59)	0.98***	0.75***	N/A
AAQ-II	1	254	19.48 (8.94)	0.70***	0.59***	0.73***
AAQ-II	2	110	21.71 (10.25)	0.85***	0.55***	0.84***
SCS	1	254	78.05 (20.70)	-0.71***	-0.77***	-0.80***
SCS-SF	2	110	38.33 (9.51)	-0.87***	-0.51***	-0.87***
FSCRS Reassured Self	1	254	21.76 (8.76)	-0.65***	-0.66***	-0.72***
FSCRS Hated Self	1	254	36.79 (15.12)	0.69***	0.62***	0.73***
FSCRS Inadequate Self	1	254	24.91 (9.99)	0.66***	0.60***	0.71***
HADS-Depression	1	254	3.53 (3.41)	0.58***	0.44***	0.59***
HADS-Anxiety	1	254	7.58 (4.40)	0.51***	0.53***	0.57***
GHQ-12	2	110	12.63 (6.58)	0.72***	0.43***	0.70***

AAQ-II, Acceptance and Action Questionnaire; SCS, Self-compassion Scale; FSCRS, Forms of Self-Criticizing/Attacking & Self-Reassuring Scale; HADS, Hospital Anxiety and Depression Scale, GHQ-12, General Health Questionnaire – 12 items.

Table 7. Hierarchical regression analysis predicting variance in HADS-Depression

Variable	Cumulative		Simultaneous	
	R^2 Change	F-change	β	p
Step 1				
Gender	0.011	F(2,231)=1.241	-0.101	0.047
Age			-0.014	0.784
Step 2				
Inadequate-self FSCRS	0.321	F(1,230)=110.441***	0.286	0.000
Step 3				
FoReST	0.098	F(1,229)=39.231***	0.431	0.000

Table 8. Hierarchical regression analysis predicting variance in HADS-Anxiety

Variable	Cumulative		Simultaneous	
	R^2 Change	F-change	β	p
Step 1				
Gender	0.054	F(2,232)=6.61**	-0.057	0.243
Age			-0.126	0.010
Step 2				
Inadequate-self FSCRS	0.386	F(1,231)=159.07***	0.465	<0.001
Step 3				
FoReST	0.033	F(1,230)=14.36***	0.249	<0.001

INSTRUCTIONS: **Below you will find a list of statements. Please rate how true each statement is for you by circling a number next to it. Use the scale below to make your choice.**

1	2	3	4	5	6	7
Never true	Very seldom true	Seldom true	Sometimes true	Frequently true	Almost always true	Always true

When I have a critical thought about myself….

1.	.... I act in a way that makes life more difficult for me	1	2	3	4	5	6	7
2.	…. It gets me so down that I don't act in the way I should	1	2	3	4	5	6	7
3.	…. I can let the feelings it creates pass from my awareness in their own time*	1	2	3	4	5	6	7
4.	…. I don't try as hard	1	2	3	4	5	6	7
5.	…. I waste more of my time	1	2	3	4	5	6	7
6.	…. I can let it pass from my awareness in its own time *	1	2	3	4	5	6	7
7.	…. I don't treat myself the way I would like	1	2	3	4	5	6	7
8.	…. I notice it without getting too caught up in it *	1	2	3	4	5	6	7
9.	…. I don't treat others the way I would like	1	2	3	4	5	6	7

*denotes items that should be negatively scored

Figure 1 Forms of Responding to Self-critical Thoughts Scale (FoReST)