Cognition and Neurosciences

Psychometric properties and validation of the Swedish Five Facet Mindfulness Questionnaire in a clinical and non-clinical sample among meditators and non-meditators

JOSEFINE L. LILJA[1,2] (iD) FREDRIK FALKENSTRÖM,[3] (iD) CLARA ZELLEROTH,[4] EMMA JACOBSON,[1] STINA RISBERG,[5] LINNEA NISSLING[1,2] (iD) and SANDRA WEINELAND[1,2] (iD)

[1]Department of Psychology, University of Gothenburg, Gothenburg, Sweden
[2]R&D Primary Health Care, Västra Götaland, Sweden
[3]Department of Behavioral Sciences and Learning, Linköping University, Sweden
[4]Gothenburg, Sweden
[5]Bäckefors, Sweden

Over a period of 15 years several attempts to conceptualize mindfulness have been presented and revised, but there is still no clear or agreed-upon definition. The use of mindfulness-based interventions has increased in clinical and research settings the last couple of years, including in Sweden. As a clinician it is crucial to know if a treatment works through the theoretically postulated mechanisms of change. Mindfulness is a concept that is difficult to measure. The overall aim of the current project was to examine the psychometric properties of the Swedish version of the Five Facet Mindfulness Questionnaire (FFMQ_SWE) using three different studies. To test the construct validity of the FFMQ_SWE a hierarchal confirmatory factor analysis was performed in a meditating non-clinical sample, to examine if all the five facets would load on an overall mindfulness construct. Psychometric properties of the instrument were examined in a non-clinical and a clinical sample, and discriminative relationships with other variables were analysed. The convergent validity was examined by analysing the correlations between FFMQ_SWE and Hospital Anxiety and Depression Scale, Sense of Coherence and Difficulties in Emotion Regulation Scale. Test-retest reliability was tested by distributing FFMQ_SWE at two occasions. The hierarchal confirmatory factor analysis showed good fit in a population of meditators. The FFMQ_SWE showed good convergent validity and test-retest reliability in both clinical and non-clinical populations. In sum, the Swedish version of the FFMQ showed good psychometric properties and can be a useful instrument as an evaluation of treatment effects in both health care settings and research settings.

Key words: Mindfulness, meditators, clinical population, psychometric properties, FFMQ.

Josefine Lilja, R&D Primary Health Care, Västra Götaland, Sweden. Tel: +46-769 40 29 69; e-mail: josefine.lilja@vgregion.se

INTRODUCTION

The use of mindfulness-based interventions (MBIs) has increased in clinical and research settings the past decades, including in Sweden. A recent study conducted in the Swedish primary care showed that Mindfulness-Based Cognitive Therapy (MBCT) can be effective as a preventive intervention for patients in primary care with recurrent depression (Lilja, Zelleroth, Axberg & Norlander, 2016). To date, MBIs are foremost recommended for people with recurrent episodes of depression (Segal, Williams & Teasdale, 2002), stress (Kabat-Zinn, 2013) and pain (Hayes, 2002).

As a clinician it is crucial to know if a treatment works through the theoretically postulated mechanisms of change. Gu, Strauss, Bond and Cavanaugh (2015) presented a meta-analysis of mediation studies MBIs. They found strong consistent evidence of cognitive and emotional reactivity, moderate and consistent evidence for mindfulness, rumination, and worry, and preliminary but insufficient evidence for self-compassion and psychological flexibility as mechanisms underlying MBIs. In order to study mechanisms of change, we need reliable and valid instruments for measuring the mechanisms we want to study.

Over a period of 15 years, several attempts to conceptualize mindfulness have been presented and revised, but there is still no clear or agreed-upon definition (Park, Reilly-Spong & Gross, 2013). In research, mindfulness has been conceptualized both as a state of practice in meditation (e.g., Lau, Bishop, Segal *et al.*, 2006) and as a trait, a predisposition to be mindful in daily life (e.g., Baer, Smith, Hopkins, Krietemeyer & Toney, 2006). Without intervention, trait mindfulness appears to be stable over time (e.g., Brown & Ryan, 2003). However, several studies have found that MBIs generally increase trait mindfulness, and that such increased trait mindfulness contributes to psychological health (Carmody, Reed, Kristeller & Merriman, 2008; Shahar, Britton, Sbarra, Figueredo & Bootzin, 2010; Shapiro, Oman, Thoresen, Plante & Flinders, 2008). Heightening of state mindfulness in meditation practice over time has been shown to lead to increases in trait mindfulness (Kiken, Garland, Bluth, Palsson & Gaylord, 2015).

Mindfulness is a concept that is difficult to measure. However, to measure inter- and intra-individual differences in mindfulness, several research groups have developed instruments in the form of self-assessment questionnaires. These include the Mindful Attention Awareness Scale (MAAS; Brown & Ryan, 2003), the Kentucky Inventory of Mindfulness Skills (KIMS; Baer, Smith & Allen, 2004), the Freiburg Mindfulness Inventory (FMI; Buchheld, Grossman & Walach, 2001), the Toronto Mindfulness

Scale (TMS; Lau *et al*., 2006), the Cognitive and Affective Mindfulness Scale, revised (CAMS-R; Feldman, Hayes, Kumar, Greeson & Laurenceau, 2007), the Southampton Mindfulness Questionnaire (SMQ; Chadwick, Hember, Symes, Peters, Kuipers & Dagnan, 2008), the Philadelphia Mindfulness Scale (PHLMS; Cardaciotto, Herbert, Forman, Moitra & Farrow, 2008), the Mindfulness/Mindlessness Scale (MMS; Haigh, Moore, Kashdan & Fresco, 2011), and the Experiences Questionnaire (EQ; Fresco, Moore, van Dulmen *et al*., 2007). One of the well-studied and reliable measures of mindfulness is the Five Facet Mindfulness Questionnaire (FFMQ) (Aguado, Luciano, Cebolla, Serrano-Blanco, Soler & García-Campayo, 2015; Baer *et al*., 2006; Lilja, Frodi-Lundgren, Johansson *et al*., 2011).

The FFMQ was originally developed from an exploratory factor analysis of pooled items from five other mindfulness questionnaires. This analysis yielded a five-factor solution (Baer *et al*., 2006). The factors/facets reflect: (1) Observing (Observe, for short); the act of noticing or attending to internal and external experiences such as sights, sounds, sensations, cognitions, and emotions; (2) Describing (Describe, for short); labelling internal experiences with words; (3) Acting with awareness (Actaware, for short); attending to one's activities of the moment, in contrast to behaving automatically while attention is focused elsewhere; (4) Non-judging of inner experiences (Non-judge, for short); taking a non-evaluative stance towards thoughts and feelings, and (5) Non-reactivity to inner experiences (Non-react, for short); allowing thoughts and feelings to come and go without getting caught up in or carried away by them (Baer *et al*., 2006).

A subsequent confirmatory factor analysis (CFA) tested this model. However, in a sample with little mindfulness experience, the FFMQ Observe facet did not load on the overall mindfulness factor (Baer *et al*., 2006). Also, items of the Observe facet did not correlate as expected with items of the Non-judge facet. When the items comprising the Observe facet were removed and the factor analysis re-run, a hierarchical four factor solution showed a good fit to the data. This result was also replicated in a Swedish study by Lilja *et al*. (2011). In this study Lilja *et al*. (2011) also developed a Swedish version of the FFMQ, which was named FFMQ_SWE. Baer *et al*. (2006) hypothesized that individuals without meditation experience may typically have more difficulties attending to inner experiences without evaluating or judging them. To test this, the factor analysis was repeated in a sample of experienced meditators (Baer *et al*., 2008). In this sample, the Observe facet loaded on the overall mindfulness factor and a five-factor structure fit the data.

To further test the idea that the ability to observe one's inner experiences is more elaborated among individuals with more meditation experience, Lilja, Josefsson, Lundh, and Falkenström (2013) approached this problem by a person-oriented approach, focusing on patterns of FFMQ facets, using cluster analysis. The analysis showed that meditators were overrepresented in four clusters, all of which had higher than average observing scores and meditators were underrepresented in three clusters, all of which had lower than average observing scores. The study also found that there seems to be a complex relationship between the ability to observe one's experiences and the ability to stay non-judgmental towards them. The results indicate that even among individuals with high overall levels of mindfulness, the abilities to observe experiences, and at the same time, keep a non-judgmental

attitude do not always go hand in hand. This study and other person-oriented studies support the assumption that Observe is an essential dimension of mindfulness, that its relationship towards the other facets needs further research and that mindfulness is a multidimensional skill that can be cultivated by meditation practice (Bravo, Boothe & Pearson, 2016; Lilja *et al*., 2013; Pearson, Lawless, Brown & Bravo, 2015).

FFMQ has previously been validated in populations with little meditative experience, that is, clinical and non-clinical populations (Baer *et al*., 2006; Cebolla, García-Palacios, Soler, Guillen, Baños & Botella, 2012; Dundas, Vollestad, Binder & Sivertsen, 2013). Jensen, Krogh, Westphael and Hjordt (2019) have shown that in a non-clinical population, the FFMQ facet scores entail an internally-consistent four-dimensional construct (all except Observe) reflecting long-term reliable dispositions that can predict mental health. Some studies (Dundas, Vollestad, Binder & Sivertsen, 2013; Van Dam, Hobkirk, Danoff-Burg & Earleywine, 2012) have proposed that the FFMQ should mainly be interpreted using scale scores, and that it can be recommended in studies seeking to differentiate between different aspects of mindfulness and how these may change overtime.

To conclude, there is a need to further investigate the FFMQ and its factor structure in clinical populations and in samples with a broad variety of meditation experience. To our knowledge there have been no studies in which the validity and reliability of the FFMQ_SWE has been tested in clinical outpatient samples. Psychometric analyses should be of interests to perform on clinical populations, since the FFMQ is designed to be used in clinical settings were patients take part in MBIs. The overall aim of the current project is to contribute to previous research on FFMQ_SWE by examining both clinical and non-clinical samples, meditators and non-meditators, as well as further investigation of psychometric properties of theFFMQ_SWE. For this purpose, three studies were carried out: one to test the construct validity of FFMQ_SWE in a Swedish sample of meditators; second, to test the convergent validity of FFMQ_SWE in a clinical and non-clinical sample; and third, to examine the test-retest reliability of the FFMQ_SWE in a clinical and a non-clinical sample.

All patients were informed of the study and consented to be a part of the study. The study was approved by The Regional Ethics Committee in Gothenburg (D.nr: 316-08).

STUDY 1

As described above, the FFMQ has been developed using different statistical methods such as exploratory factor analysis to test the construct validity. Construct validity refers to the extent to which a test captures a specific theoretical construct. It is a demonstration that the phenomenon being measured exists in the way it is theoretically described, and theoretical models need to be continuously evaluated, modified and adapted (Clark-Carter, 2010; Pett, Lackey & Sullivan, 2003). The purpose of the first study was to investigate and replicate the hierarchical five-factor structure that was shown by Baer *et al*. (2008) in a sample with meditation experience in a Swedish population, using CFA. The research questions in Study 1 concerning construct validity were: (1) is the Observe facet stronger related to the global mindfulness construct in a sample of meditators, than the other facets in the

FFMQ_SWE?; and (2) do the five facets in FFMQ_SWE all load on an overall mindfulness construct (i.e. a second-order factor analysis)?

Method

Procedures and participants. The participants represented a combination of several different samples, recruited as part of three other studies on mindfulness: (1) *The Gothenburg sample.* This sample included 498 individuals (including both university students and participants from meditation centers, as well as others). The participants from the meditation centers and those who reported "a lot" and "fair amount" of meditation experience were combined into one subgroup of meditators ($n = 85$). (2) *The yoga sample.* Several yoga centers in Malmö, Sweden, were contacted, asking if they wanted to participate in a study on meditating individuals' experiences. Five different centers (Classic Yoga, Natha Yoga, Yoga Kendra, Bikram Yoga, and Manfrinato Yoga) responded positively to this request, and a research assistant visited these centers in connection with yoga classes, asking participants to fill out a questionnaire that included the FFMQ and some questions about age, gender and experiences of meditation. In total 153 individuals from the yoga centers filled out the questionnaire. (3) *The Vipassana sample.* This was a sample of 85 individuals from a Vipassana center in Stockholm, who participated in a quasi-experimental study of changes in self-reported mindfulness after participating in a silent meditation retreat arranged by this center (Falkenström, 2010). The present study used data from the participants as well as the comparison group before the retreat (all participants had taken part in at least one earlier silent retreat).

Of these 736 participants, there was full data on the FFMQ for 734 individuals. All participants answered questions about their experience of meditation. Of these, 452 responded that they had no experience of meditation, or reported either that they had "little" experience of meditation (in the Gothenburg sample) or that they had less than one year's experience of meditation (in the yoga and Vipassana samples), and therefore fell outside the comparison groups and were excluded from the sample. Of the remaining participants, 282 responded that they had at least one year's experience of meditation (in the yoga sample, and the Vipassana sample), or that they had at least "fair amount" or "extensive" experience of meditation (in the Gothenburg sample). Among the participants, 65.6% ($n = 185$) were women and 34.0% ($n = 96$) were men, one individual did not state their gender. All participants from the yoga sample (Sample 2) and the Vipassana sample (Sample 3), were currently practicing meditation at the time of the study; we did not, however, have information about current meditation practice from the meditator subsample of Sample 1. Age distribution of the 282 participants was as follows: < 21 years, $n = 6$; 21–24 years, $n = 11$; 25–30 years, $n = 37$; 31–40 years, $n = 93$; 41–50 years, $n = 63$; >50 years, $n = 70$. Two individuals did not state their age.

Instrument

Five Facet Mindfulness Questionnaire (Baer et al., 2006) – Swedish version of the FFMQ (FFMQ_SWE) (Lilja *et al.*, 2011). The instrument has a global scale and five subscales referring to five facets of people's general tendency to be mindful in daily life. The facets are named Observe, Describe, Actaware, Non-judge and Non-react (Baer *et al.*, 2006). Items are rated on a five-point Likert-type scale ranging from 1 (never or rarely true) to 5 (very often or always true). All five facets and the Global Scale (i.e. overall mean of all items) of the FFMQ_SWE shows good internal consistency: Observe $\alpha = 0.83$, Describe $\alpha = 0.91$, Actaware $\alpha = 0.87$, Non-judge $\alpha = 0.87$ and Non-react $\alpha = 0.75$, FFMQ_SWE Global Scale $\alpha = 0.81$ (Lilja *et al.*, 2011). Mindfulness facets have been shown to be differentially correlated with several other constructs and to have incremental validity in

the prediction of psychological symptoms (Baer *et al.*, 2006). In the present study, reliability was good for FFMQ_SWE Global Scale $\alpha = 0.87$, Non-react $\alpha = 0.82$, Observe $\alpha = 0.75$, Actaware $\alpha = 0.83$, Describe $\alpha = 0.83$, and Non-judge $\alpha = 0.83$.

Data analysis. The CFA was carried out using Mplus 8 version 1.5 (Muthén & Muthén, 1998–2017), and the input raw data were stored in PASW 23. CFA was used to statistically test a hypothesized model (i.e., whether the Swedish data confirm the theoretical model of mindfulness) in a population of meditators. To replace missing values, we calculated the mean value from the respondent's answer in the facet. The questionnaires had few missing answers (internal missing less than 1%). Several models were tested to see which one had the best model fit. The model Chi-square test, although performing best in simulations, tends to reject models with large N and/or large degrees of freedom in real data. Presumably this is because real data contain many small residual covariances due to for example wording effects that, although non-random, do not necessarily invalidate the hypothesized model (Muthén & Asparouhov, 2012). In the present study, we therefore used recommended cut-offs for fit indices; the comparative fit index (CFI), the root mean squared error of approximation (RMSEA), and the standardized root mean square residual (SRMR; Kline, 2005). CFI values greater than 0.90 are considered to indicate a reasonably good fit between the data and the model. For the RMSEA, a value of 0.05 is considered a close fit and values up to 0.08 acceptable fit (Kline, 2005). In addition, a 90% confidence interval (CI) around the RMSEA value is reported. The modification index (MI) was used to look for potential improvement of the model. The MI is used to measure the reduction of the chi-square when a specific change in the model has been applied (Byrne, 2010).

Results

Confirmatory factor analysis of the FFMQ_SWE on meditators. First, a five-factor model was estimated using maximum likelihood with robust standard errors. In this model, all five factors were allowed to correlate freely (model 1, see Fig. 1). This model yielded reasonable model fit (χ^2 (367) = 613.06, $p < 0.001$; RMSEA = 0.049, 95% CI 0.042, 0.055, probability that RMSEA <0.05 = 0.61; CFI = 0.90; SRMR = 0.064). Standardized factor loadings were all> 0.43.[1] All factor correlations were statistically significant except Non-judge with Observe ($p = 0.47$). To test the hypothesis that all factors are part of an over-arching mindfulness construct, a second-order factor model (model 2, Fig. 1) was estimated in which all five factors were used as indicators of an overall mindfulness factor. This model also fit reasonably well (χ^2 (372) = 633.31, RMSEA = 0.050, 95% CI 0.043, 0.056, probability that RMSEA <0.05 = 0.50, CFI = 0.89, SRMR = 0.071). The deterioration in model fit due to the additional constraints put on the model was, however, statistically significant (Satorra Bentler scaled $\Delta\chi^2$ (5) = 18.94, $p = 0.002$), meaning that model 2 fit significantly worse than model 1. Possible reasons for the failure of the second-order model were explored. All factors loaded significantly and reasonably strong (> 0.42) on the second order factor. However, modification indices

showed that a residual correlation between Non-judge and Observe would improve model fit. The model was re-run with this residual correlation added (model 3, Fig. 1). The model fit (χ^2 (371) = 616.34, RMSEA = 0.048, 95% CI 0.042, 0.055, probability that RMSEA <0.05 = 0.64, CFI = 0.90, SRMR = 0.066) was significantly better than the unmodified second-order model (Satorra Bentler scaled $\Delta\chi^2$ (1) = 21.54, $p < 0.001$) and non-significantly worse than the saturated model (Satorra Bentler scaled $\Delta\chi^2$ (4) = 3.87, $p = 0.42$). Factor loadings on the second-order factor were now higher for both Non-judge (0.65 vs 0.60) and Observe (0.52 vs 0.44). The smallest

standardized factor loading was for the Describe facet (0.42). The residual correlation between Non-judge and Observe was negative (standardized estimate = -0.43), indicating that after taking the common (second-order) mindfulness factor into account these two facets correlated negatively.

STUDY 2

The purpose of Study 2 was to examine the validity of FFMQ_SWE in a clinical and a non-clinical population concerning the relationship between mindfulness, emotional
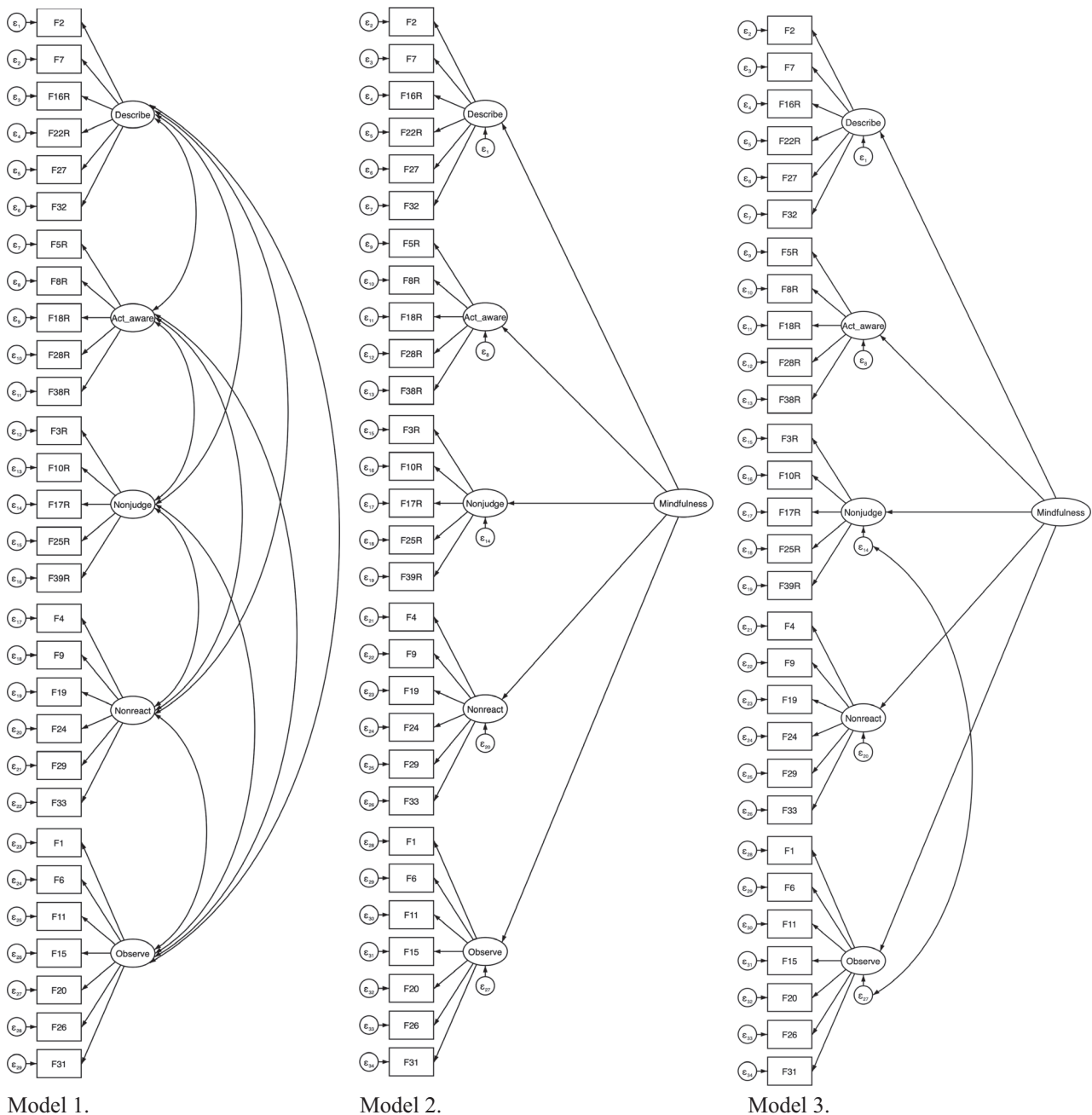

Model 1.   Model 2.   Model 3.

Fig. 1. Confirmatory factor analysis models tested. Model 1 is a standard correlated factors model, while models 2 and 3 are second-order models in which first-order factors load on a overall mindfulness factor. Model 3 is a post hoc modification based on inspection of Modification Indices, which adds a residual covariance between Non-judge and Observe

regulation and mental symptoms, as well as the relationship between mindfulness and sense of coherence. Convergent validity is a form of construct validity that aims to test if two, or more, measures that are supposed to be measuring the same construct are related (Clark-Carter, 2010). The research questions in Study 2 concerning convergent validity were: (1) is the FFMQ_SWE negatively correlated with depressive and anxiety symptoms and positively correlated with sense of coherence in a Swedish non-clinical sample?; (2) is FFMQ_SWE negatively correlated with depressive and anxiety symptoms and emotional regulation, and positively correlated with sense of coherence, in a Swedish clinical sample?; and (3) is there a significant difference in levels of mindfulness between a clinical and non-clinical population?

Method

Participants in the non-clinical sample. This was a heterogeneous sample (including both university students and participants from meditation centers, as well as others) who were recruited as part of a reliability study of the Swedish version of the FFMQ carried out at the Department of Psychology, Gothenburg University (Lilja *et al.*, 2011). The non-clinical data were collected from Swedish university students, health care practitioners, teachers at a Swedish university, and the general population (see Table 1). The sample was a convenience sample and the participants were contacted via mail or personal contact. No compensation for participation was offered. The sample also included 22 mindfulness meditators, who were recruited from different Buddhist meditation centers in Gothenburg, Sweden. The total number of participants was 498 (296 women and 197 men; five participants did not state their gender); however, three individuals left the FFMQ blank, resulting in a sample size of 495. Age distribution of participants was as follows: < 21 years, $n = 3$; 21–24 years, $n = 178$; 25–30 years, $n = 63$; 31–40 years, $n = 65$; 41–50 years, $n = 53$; >50 years, $n = 61$. Two individuals did not state their age. Participants reported the following meditation experience: 288 "none at all" (58%), 123 "a little" (25%), 67 "fair amount" (14%), and 17 "a lot" (3.4%). In one questionnaire, meditation experience was not stated. The participants from the meditation centers and those who reported "a lot" and "fair amount" of meditation experience were combined into one subgroup of meditators ($n = 85$). Descriptive statistics for the non-clinical sample on the instruments were Hospital Anxiety and Depression Scale (HADS-Anxiety), M = 6.32 (SD = 3.57), (HADS-Depression), M = 3.22 (SD = 2.43) and Sense of Coherence (SOC), M = 64.41 (SD = 10.52).

Table 1. *Demographic characteristics of the sample*

Sample	Numbers of participants
Students from Halmstad College and Gothenburg University	306
Practitioners at a Maternal and Child Health Clinic	31
Practitioners of cognitive and behavioral therapy at Child and Adolescent Psychiatry	27
Practitioners at a Psychiatric Child Clinic	24
Practitioners of cognitive and behavioral psychotherapy	19
Teachers at Halmstad Elementary School	18
Health Clinic workers	24
Acquaintances	27
Mindfulness Meditators	22
Total	498

Participants in the clinical sample. Participants were recruited from two adult psychiatric outpatient clinics ($n = 48$), and three primary health care centers ($n = 109$) in western part of Sweden. Participants were patients, seeing a psychologist for their mental health issues, at one of these clinics between January 2010 and May 2018. Exclusion criteria for participation were lack of language skills in Swedish, diagnosed mental retardation, ongoing psychosis, ongoing full-time care or homelessness. In the study a total of 157 people participated, of which 103 women and 54 men. Age distribution of participants was as follows: < 21 years, $n = 5$; 21–24 years, $n = 11$; 25–30 years, $n = 23$; 31–40 years, $n = 41$; 41–50 years, $n = 34$ and 43 were older than 50 years. A total of $n = 157$ in the clinical group was used for comparisons with the non-clinical sample ($n = 495$). The participants from the adult psychiatric outpatient clinics ($n = 48$), and from two primary health care centers ($n = 26$) were combined into one sub-clinical sample where we validated the relationship between FFMQ_SWE, HADS, SOC and Difficulties in Emotion Regulation Scale (DERS), $n = 74$. Descriptive statistics for the clinical sample on the instruments were HADS-A, M = 11.97 (SD = 4.02), HADS-D, M = 7.5 (SD = 4.4), SOC, M = 46.35 (SD = 13.1) and DERS M = 49.7 (SD = 16.64).

Instruments

Five Facet Mindfulness Questionnaire (Baer et al., 2006) – Swedish version of the FFMQ (FFMQ_SWE) (Lilja *et al.*, 2011). The questionnaire is described in Study 1. In the present study, reliability was good for FFMQ_SWE Global Scale α = 0.87, Non-react α = 0.79, Observe α = 0.76, Actaware α = 0.82, Describe α = 0.85, and Non-judge α = 0.84.

Hospital Anxiety and Depression Scale (HADS). The HADS was used to measure mental symptoms (Lisspers, Nygren & Soederman, 1997; Zigmond & Snaith, 1983). It consists of 14 statements (7 on depression and 7 on anxiety) with four response alternatives (0 to 3). The HADS has been shown to be a reliable and valid instrument for the detection of anxiety and depression in individuals from 16 to 65 years of age (Zigmond & Snaith, 1983). Lisspers *et al.* (1997) tested HADS on a Swedish population 30–59 years of age (n = 624) and found that the mean on HADS-A was 4.27 (SD = 3.58) for men and 4.76 (SD = 3.83) for women, and on HADS-D, 4.27 (SD = 3.57) for men and 3.76 (SD = 3.36) for women. Its reliability was shown by Hermann (1997) with Cronbach's alpha on HADS-A at 0.80 and on HADS-D at 0.81. The maximum score on each subscale is 21, and 11 points is the cut-off level for a diagnosis of anxiety or depression. Values of 0–6 indicate no or a normal level of anxiety or depression (Zigmond & Snaith, 1983). Cronbach's alphas in this study were 0.84 on HADS-A and 0.76 on HADS-D.

Sense of Coherence 13 (SOC-13). SOC-13 was developed by Antonovsky (1987) to measure the sense of coherence, which consists of perceived comprehensiveness, manageability and meaningfulness. Eriksson and Lindström (2005) argue that SOC-13 is strongly related to mental health and is an important contributor for the development and maintenance of people's general health. SOC-13 consists of 13 statements that capture the experience of sense of coherence. The items are answered on a 1–7 Likert scale and higher scores indicate a stronger sense of coherence (Antonovsky, 1987). SOC-13 has good internal validity with Cronbach's alpha at 0.89 (Olsson, Gassne & Hansson, 2009). In a systematic review of 127 studies using SOC-13,

Eriksson and Lindström (2005) found an internal validity with Cronbach's alpha at between 0.70 and 0.92. The average score for SOC-13 vary between 35.39 (SD = 0.10) and 77.60 (SD = 13.80). The concept validity of SOC-13 is weaker than the internal validity. Cronbach alpha in this study was 0.88 on SOC-13 Total.

Difficulties in Emotion Regulation Scale - short version (DERS-16). The scale is a self-assessment questionnaire that measures the ability to regulate emotions. DERS-16 has been developed and validated in three studies in Sweden and the United States (Bjureberg, Ljótsson, Tull *et al.*, 2015). DERS-16 consists of 16 items, where the statements are answered by a 1–5 Likert scale. DERS-16 consists of five dimensions; non-acceptance, clarity, goals, impulse and strategies. Scores on DERS range between 16 and 80, higher scores indicating more difficulties with emotional dysregulation. DERS-16 has good psychometric properties, with good validity and excellent internal reliability with Cronbach alpha of 0.92 and good test-retest reliability (Bjureberg *et al.*, 2015). Cronbach alpha in this study was 0.96 on DERS-16 Total.

Data analysis. To examine the validity of FFMQ_SWE, HADS, SOC and DERS in a clinical population and FFMQ_SWE, HAD and SOC in a non-clinical population, Pearson product-moment correlation coefficient was used. We used a t-test to examine if there were any differences between the clinical and non-clinical population in level of mindfulness. In addition, a measurement invariance test was done to examine the factorial stability across clinical and non-clinical samples (Cheung & Rensvold, 2002).

Results

Validity of the FFMQ_SWE in a non-clinical sample. A Pearson product-moment correlation coefficient was computed to assess the relationship between FFMQ_SWE, HADS and SOC. There was a significant correlation between HADS and FFMQ_SWE, $r = -0.502$, $p < 0.001$. There was also a significant correlation between SOC and FFMQ_SWE, $r = 582$, $p < 0.001$ (see Table 2).

Validity of the FFMQ_SWE in a clinical sample. A Pearson product-moment correlation coefficient was computed to assess the relationships between FFMQ_SWE, HADS, SOC and DERS. There was a significant correlation between HADS and FFMQ_SWE, $r = -0.556$, $p < 0.001$. There was also a

Table 2. *Convergent validity between FFMQ_SWE, HADS and SOC in a non-clinical sample, n = 495*

FFMQ_SWE Facets	HAD-A	HAD-D	HADS	SOC
Non-react	−0.345[**]	−0.156[**]	−0.316[**]	0.305[**]
Observe	−0.026	−0.092[*]	−0.062	0.086
Actaware	−0.374[**]	−0.268[**]	−0.390[**]	**0.460**[**]
Describe	−0.208[**]	−0.224[**]	−0.252[**]	0.323[**]
Non-judge	−0.519[**]	−0.247[**]	−0.481[**]	0.537[**]
Global Scale	−0.491[**]	−0.333[**]	−0.502[**]	0.582[**]

Note: *p < 0.05, **p < 0.01.

significant correlation between SOC and FFMQ_SWE, $r = 0.679$, $p < 0.001$ (see Table 3). There was also a significant correlation between DERS and FFMQ_SWE, $r = -0.678$, $p < 0.001$ (see Table 3).

Differences between clinical and non-clinical population in levels of mindfulness. To investigate FFMQ_SWE's usefulness in clinical populations, a t-test (5% alpha level) was conducted. The t-test showed that individuals in a normal population ($n = 495$) had significant higher levels of mindfulness than individuals in a clinical population ($n = 157$), $t(650) = -13.283$, $p < 0.001$. In the normal population the average was 3.36 ($SD = 0.40$) and in the clinical population, the average rate was 2.86 ($SD = 0.46$) on FFMQ_SWE. The non-clinical population had significantly higher levels of mindfulness than the clinical population on all the facets (p-values between 0.001 and 0.016) except Observe, $t(650) = -2.794$, $p < 0.065$.

Differences between clinical and non-clinical samples in factorial structure. An initial factor analysis on the two groups combined, showed a marginal fit according to most measurements ($\chi^2 (367) = 1014.39$, $p < 0.001$; RMSEA = 0.05, 90% CI 0.05, 0.06; CFI = 0.88, SRMR = 0.06). A measurement invariance analysis comparing the clinical and non-clinical samples in terms of factor loadings and intercepts was performed. Results showed that metric invariance held, that is, factor loadings did not differ significantly between the clinical and non-clinical samples ($\Delta\chi^2 (24) = 25.80$, $p = 0.36$). However, scalar invariance did not hold, meaning that intercepts differed significantly between the two samples ($\Delta\chi^2 (24) = 122.42$, $p < 0.001$). The findings were the same even if the more liberal delta-CFI (<0.01) criterion was used.

Exploration of relation between Observe and Non-judge in the prediction of psychological distress/well-being. Due to the findings of Study 1, in which we found a residual correlation between Observe and Non-judge after controlling for the overall mindfulness factor, we also wanted to explore whether these two facets would interact in predicting psychological distress or well-being. Theoretically, it seemed plausible that persons with high scores on Observe, together with low scores on Non-judge, would be more likely to experience psychological distress and less likely to have a strong sense of coherence than person with high scores on both these variables. We therefore ran two post hoc regression models, one with main effects of Observe and Non-judge, and another that also included the interaction term between Observe and Non-judge (i.e., Observe × Non-judge). Results showed that the interaction effect was non-significant for all measures in both the samples. However, in the main effects model, Observe significantly predicted lower distress and stronger sense of coherence in the non-clinical sample (HADS: Observe $b = -0.96$, SE = 0.30, $\beta = -0.13$, $p < 0.004$, Non-judge $b = -3.12$, SE = 0.25, $\beta = -0.50$, $p < 0.001$; SOC: Observe $b = 2.45$, SE = 0.59, $\beta = 0.16$, $p < 0.001$; Non-judge $b = 7.19$, SE = 0.49, $\beta = 0.56$, $p < 0.001$). Thus, holding Non-judge constant, being better able to Observe one's moment-to-moment experiences was associated with better health in the non-clinical sample. This was, however, not the case in the clinical sample.

Table 3. *Convergent validity between, FFMQ_SWE, HADS, SOC and DERS-16 in a clinical population, n = 74*

FFMQ_SWE Facets	HAD-A	HAD-D	HADS	SOC	DERS-16
Non-react	−0.562[**]	−0.388[**]	−0.536[**]	0.606[**]	−0.705[**]
Observe	−0.039	−0.169	−0.110	0.143	−0.099
Actaware	−0.563[**]	−0.416[**]	−0.546[**]	**0.657**[**]	−0.588[**]
Describe	−0.228	−0.221	−0.239[*]	0.297[*]	−0.209
Non-judge	−0.396[**]	−0.209	−0.356[**]	0.494[**]	−0.591[**]
Global Scale	−0.546[**]	−0.449[**]	−0.556[**]	0.679[**]	−0.678[**]

Note: *$p < 0.05$, **$p < 0.01$.

STUDY 3

One goal of research is to establish laws and principles that have a certain general applicability. In an experiment with multiple measurements, the measurement tool needs to consistently reproduce the same result, providing that all other variables remain the same. An instrument with good test re-test reliability strengthens the internal validity and ensures that the measurements are both representative and stable over time (Clark-Carter, 2010). The purpose of Study 3 was to evaluate the test-retest reliability of FFMQ_SWE by examining how strong the correlation was between occasion 1 and 2 in a Swedish non-clinical and clinical sample. The research questions in Study 3 concerning test-retest reliability were: (1) how strong is the correlation between FFMQ_SWE on occasion 1 and 2 in Swedish non-clinical and clinical samples?; and (2) is there a difference in the test-retest reliability between the clinical and non-clinical sample?

Method

Participants. The clinical sample ($n = 31$) came from a primary health care centre in the Västra Götalandsregionen. The non-clinical sample were 29 randomly selected individuals from the sample of Study 2 that agreed to answer the FFMQ_SWE three weeks after the first occasion.

Instruments

Five Facet Mindfulness Questionnaire (Baer et al., 2006) – Swedish version (FFMQ_SWE) (Lilja et al., 2011). The questionnaire is described in Study 1.

Data analysis. Test-retest reliability was analysed with Pearson's product-moment correlation coefficient. Internal consistency was analysed using Cronbach's coefficient alpha.

Table 4. *Test-retest correlations among the clinical sample, n = 31*

FFMQ domains	Evaluated	Test-retest values	M(SD) measure1	M(SD) measure2
Non-react		0.834[**]	3.333 (0.519)	3.446 (0.527)
Observe		0.722[**]	3.448 (0.659)	3.575 (0.582)
Actaware		0.356[*]	3.290 (0.518)	3.380 (0.559)
Describe		0.705[**]	3.963 (0.401)	3.898 (0.603)
Non-judge		0.646[**]	3.806 (0.664)	3.967 (0.756)
Global Scale		0.509[**]	3.564 (0.331)	3.649 (0.317)

Note: *$p < 0.05$, **$p < 0.01$.

Table 5. *Test-retest correlations among the non-clinical sample, n = 29*

FFMQ domains	Evaluated	Test-retest values	M(SD) measure1	M(SD) measure2
Non-react		0.359	2.44 (0.63)	2.58 (63)
Observe		0.751[**]	2.67 (0.71)	2.63 (0.75)
Actaware		0.706[**]	3.01 (0.94)	2.93 (0.80)
Describe		0.830[**]	3.02 (0.68)	2.91 (74)
Non-judge		0.595[**]	2.82 (0.74)	2.57 (0.80)
Global scale		0.778[**]	2.77 (0.36)	2.73 (0.38)

Note: *$p < 0.05$, **$p < 0.01$.

Results

Test-retest reliability and internal consistency in non-clinical and clinical samples. Descriptive statistics and test-retest-reliability coefficient for FFMQ_SWE with facets is presented in Tables 4, 5 and 6. The results showed significant test-retest correlations for the FFMQ_SWE Global Scale and all the facets in the clinical sample, $n = 31$ (see Table 4). For the non-clinical sample, the test-retest correlations were significant for the FFMQ_SWE Global Scale and all the facets, except Non-react (see Table 5).

The reliability coefficient for the FFMQ_SWE Global Scale in the non-clinical sample was considered to be good. In the clinical sample the correlation was considered to be acceptable. Cronbach's alpha for the FFMQ_SWE Global Scale was acceptable in the non-clinical sample and poor in the clinical sample (see Table 6).

DISCUSSION

Psychometrically sound instruments are needed to evaluate mechanisms of change in mindfulness treatments. The aim of the present study was to examine the psychometric properties of the FFMQ_SWE and validate the questionnaire, using three different study samples. The results show that a five-factor model fit FFMQ_SWE reasonably well for individuals with meditation experience. Our results also show that FFMQ_SWE possesses satisfactory psychometric properties in both a clinical and non-clinical population. The FFMQ_SWE shows good internal consistency and the repeated measures analysis suggests that the FFMQ_SWE is a stable measure over time. Moreover, FFMQ_SWE show a pattern of theoretically consistent correlations regarding anxiety, depression, sense of coherence and difficulties in emotional regulation.

Table 6. *Cronbach alpha FFMQ_SWE, in a non-clinical and clinical sample*

FFMQ_SWE – non-clinical sample	Cronbach alpha	Reliability Index	FFMQ_SWE - clinical sample	Cronbach alpha	Reliability Index
Non-react	0.359	Questionable	Non-react	0.834[**]	Good
Observe	0.751[**]	Good	Observe	0.722[**]	Good
Actaware	0.706[**]	Good	Actaware	0.356[*]	Questionable
Describe	0.803[**]	Good	Describe	0.705[**]	Good
Non-judge	0.595[**]	Acceptable	Non-judge	0.646[**]	Acceptable
Global Scale	0.778[**]	Good	Global Scale	0.509[**]	Acceptable

Note: *$p < 0.05$, **$p < 0.01$.

Study 1

The hierarchical model, with a second-order mindfulness construct explaining the covariances among the five facets, fits reasonably well but not as well as the less constrained model. Post hoc exploration shows that this could be explained by a negative residual covariance between Non-judge and Observe. The negative residual covariance can be interpreted as an effect of some individuals scoring high on Non-judge and low on Observe. However, this is contradicted by other individuals scoring high on both Non-judge and Observe as indicated by the strong loadings on the FFMQ_SWE Global Scale. Relating this to our research questions, it seems that among meditators, the ability to observe one's emotional state is indeed part of the overall mindfulness construct. Nevertheless, there was still some heterogeneity in this regard, indicated by the relationship between Observe and Non-judge being negative for some participants. It is possible that this was due to differences in meditation experience within the meditator sample, but unfortunately it was not possible to test this since meditation experience had been scored differently in different samples. This finding means that for some participants, close observation of internal experience is associated with negative judgements about the self, while for others it is associated with self-acceptance. The definition of mindfulness is that of self-observation in a non-judging manner, namely, the combination of these two facets. This means that the two facets do not contribute to overall mindfulness independently, but in interaction with each other.

All existing definitions of mindfulness refer to the activity of observing (paying attention to) one's external and internal experiences as part of the very core of mindfulness. Yet in Baer et al.'s (2006) study and Lilja et al.'s (2011) study, the Observe facet failed to fit in as a component of the overall mindfulness construct in a non-meditating population. The results from Baer and co-workers also showed that four of the facets (all but Observe) were consistently related in expected ways to a variety of other variables. In a second study by Baer and co-workers (2008), data were collected from experienced meditators, and a CFA was performed. The results confirmed a hierarchical five-factor model among individuals with meditation experience. These findings are partly consistent with our results from Study 1 and support the hypothesis that all the facets in mindfulness are part of an overall mindfulness in a meditating sample. Furthermore, Baer et al. (2008) found that the positive relationship between Observe and symptoms of well-being varied with meditation experience. Our results, in conjunction with previous studies (Aguado, Luciano, Cebolla, Serrano-Blanco,

Soler & García-Campayo, 2015; Baer et al., 2008; Jensen, Krogh, Westphael & Hjordt, 2019; Veehof, ten Klooster, Taal, Westerhof & Bohlmeijer, 2011) indicate that the ability to observe one's feelings and inner experiences, and its relation to the overall ability to be mindful, is significant only in the meditating sample and that the Observe facet should be interpreted with caution in a non-meditative population.

Study 2

The FFMQ_SWE shows convergent validity with consistent correlations regarding anxiety, depression, sense of coherence and emotional regulation in both a clinical and non-clinical population. These results are in line with Bjureberg et al. (2015) who in a validation study showed that there are strong correlations between emotional regulation and levels of mindfulness. However, the results from the current study show that the ability of both the clinical and non-clinical samples to observe one's experiences and feelings (Observe facet), are not significantly related to symptoms of anxiety and depression, nor to sense of coherence. Some theoretical points can be made from this.

The findings regarding the FFMQ facets showing consistent correlations apart from the Observe facet, can be interpreted as further evidence of acceptance as being more important than present-centered attention regarding mental health. Coffey, Hartman and Fredrickson (2010) investigated the relationship between mindfulness, emotional regulation, and well-being by executing a combination of exploratory and confirmatory factor analyses and found that acceptance of one's experience matters more for mental health than present-centered attention. The findings are in line with the results regarding the Observe facet from person-oriented studies of mindfulness (Bravo, Boothe & Pearson, 2016; Lilja, Josefsson, Lundh & Falkenström, 2013; Pearson, Lawless, Brown & Bravo, 2015). These studies suggest that mindfulness is a multidimensional concept and that the abilities that we try to measure with FFMQ are affecting each other. A possible explanation is that the qualitative aspect, *how* you observe your feelings and thoughts and what you do with these observations, for example being judgmental or reactive, affects your mental health. In line with this, our post hoc exploratory regression model shows that when acceptance was held constant statistically, present-centered observation was related to mental health in the non-clinical sample. Theoretically, this makes sense since the definition of mindfulness does not imply that the components of acceptance and observation are independent predictors of mindfulness but should be present

together. The fact that this was not the case in the clinical sample is harder to explain, but it is possible that the items in FFMQ are interpreted differently by the clinical and non-clinical populations (which the partial failure of measurement invariance also indicates).

Used accordingly, training in MBIs should give the gradual acquirement of mindfulness skills, and therefore a successive convergence between the abilities to observe *and* to keep a non-judgmental attitude. These results support the assumption that FFMQ should mainly be interpreted using the facets (Dundas, Vollestad, Binder & Sivertsen, 2013; Van Dam, Hobkirk, Danoff-Burg & Earleywine, 2012) and that a person-oriented approach (patterns between the facets) can help clinicians to tailor specific mindfulness-based practices to individuals based on their mindfulness profiles.

The analysis shows that the non-clinical sample had a higher general tendency to be mindful in daily life than the clinical sample. In that sense the FFMQ_SWE fulfills a pattern of theoretical and clinical assumptions regarding how one would expect relationships to unfold between the ability to be mindful, psychological processes and symptoms. An interesting finding is that in the clinical sample the result shows that the ability to put inner experiences into words (Describe facet) does not correlate significantly with emotional regulation. These results suggest that in clinical therapeutic work it is not only important that the patient learn to describe their feelings and thoughts, but also learn how to deal with these expressions.

Furthermore, the clinical and non-clinical participants respond to the items of the FFMQ in different ways. This should be kept in mind when using the FFMQ to compare mean values between different groups of respondents (at least clinical and non-clinical groups). Mean values can be compared within the groups, but caution should be used when comparing across groups.

Study 3

Results show that the test-retest reliability of the FFMQ_SWE Global Scale in the non-clinical sample is considered to be good. The results are in line with previous studies on test-retest reliability. Jensen, Krogh, Westphael and Hjordt (2019) state that FFMQ has adequate short-term two-weeks test-retest reliability among a non-clinic population. In this study the test-retest reliability of the FFMQ_SWE Global Scale in the clinical sample is considered to be acceptable, however the Non-react facet is not reliable regarding test-retest in the non-clinical sample. In addition, the Observe facet do not differ between the clinical and non-clinical sample. In sum, results show that the total score of FFMQ_SWE can be considered a stable measure over time and can be used in studies evaluating mindfulness interventions. However, some caution should be taken regarding some of the facets in the clinical and non-clinical sample.

Descriptive statistics for the clinical groups show that scores on HADS was generally above the cut-off levels of a Swedish population (Lisspers, Nygren & Soederman, 1997). The current study makes an important contribution by providing information about the validity and reliability of using the FFMQ_SWE in clinical settings when assessing patients for

MBIs. The results can therefore be generalized to clinical and research settings when evaluating the effects of a given MBI.

Limitations and further research

A methodological limitation of the current study is that the meditation experience had been scored in different ways in different samples. It was therefore not possible to make any deeper analysis in case this impacted the scoring of the observe facet. In future research it would be interesting to study clinical samples who receive MBIs and to use a person-oriented research approach to identify subgroups based on their mindfulness scoring. Furthermore, we would like to examine whether mindfulness interventions can impact the mindfulness profiles of clinical samples and thereby their overall psychological wellbeing. In continuum, this would allow for clinicians to interpret the answers on the FFMQ facets and more easily tailor treatment interventions. Another limitation is that the participants in Study 2 came from different clinics, the public population and it was not possible to control the age or gender of the sample. As an effect there are more young people represented in the non-clinical sample.

CONCLUSIONS AND IMPLICATIONS

FFMQ_SWE shows valid and sound psychometric properties. FFMQ_SWE is a potential questionnaire to be used as an assessment instrument, to tailor treatment interventions and to evaluate treatment outcomes, and can therefore be used in both health care settings and research settings. Results also indicate that the Observe facet has a complex relationship to the overall mindfulness construct, and that scores on this facet cannot be interpreted in isolation from other facets – in particular the Non-judge facet. Mindfulness is not about simply observing experience, and not simply acceptance, but of the combination of close observation with acceptance.

Future research should attempt to replicate findings from this study, as well as test experimentally whether mindfulness as measured by the FFMQ_SWE can be improved by treatment (e.g. MBSR or MBCT).

NOTE

[1] A method factor loading on all negatively worded items (loadings constrained to equality, correlation with other factors constrained to zero) was added to test whether the wording of the items affected the factor structure. The addition of the method factor significantly improved model fit ($p < 0.001$), but loadings on this factor were small (standardized loading = 0.22) and changes to other aspects of the model were negligible.

REFERENCES

Aguado, J., Luciano, J. V., Cebolla, A., Serrano-Blanco, A., Soler, J. & García Campayo, J. (2015). Bifactor analysis and construct validity of

the five facet mindfulness questionnaire (FFMQ) in non-clinical Spanish samples. *Frontiers in Psychology*, 6, 404. https://doi.org/10.3389/fpsyg.2015.00404.

Antonovsky, A. (1987). *Unravelling the mystery of health: How people manage stress and stay well*. San Francisco, CA: Jossey-Bass.

Baer, R. A., Smith, G. T. & Allen, K. B. (2004). Assessment of mindfulness by self-report: The Kentucky inventory of mindfulness skills. *Assessment*, 11, 191–206.

Baer, R. A., Smith, G. T., Hopkins, J., Krietemeyer, J. & Toney, L. (2006). Using self-report assessment methods to explore facets of mindfulness. *Assessment*, 13, 27–45.

Baer, R. A., Smith, G. T., Lykins, E., Button, D., Krietemeyer, J., Sauer, S. *et al.* (2008). Construct validity of the five facet mindfulness in meditating and non-meditating samples. *Assessment*, 15, 329–342.

Bjureberg, J., Ljótsson, B., Tull, M. T., Hedman, E., Sahlin, H., Lundh, L.-G. *et al.* (2015). Development and validation of a brief version of the Difficulties in Emotion Regulation Scale: The DERS-16. *Journal of Psychopathology and Behavioral Assessment*, 38, 284–296.

Bravo, A. J., Boothe, L. G. & Pearson, M. R. (2016). Getting personal with mindfulness: A latent profile analysis of mindfulness and psychological outcomes. *Mindfulness*, 7, 420–432.

Brown, K. W. & Ryan, R. M. (2003). The benefits of being present: Mindfulness and its role in psychological well-being. *Journal of Personality and Social Psychology*, 84, 822–848.

Buchheld, N., Grossman, P. & Walach, H. (2001). Measuring mindfulness in insight meditation (Vipassana) and meditation-based psychotherapy: The development of the Freiburg Mindfulness Inventory (FMI). *Journal for Meditation and Meditation Research*, 1, 11–34.

Byrne, B. M. (2010). *Structural equation modelling with AMOS: basic concepts, applications, and programming* (2nd edn). New York: Routledge.

Cardaciotto, L., Herbert, J. D., Forman, E. M., Moitra, E. & Farrow, V. (2008). The assessment of present-moment awareness and acceptance: The Philadelphia Mindfulness Scale. *Assessment*, 15, 204–223.

Carmody, J., Reed, G., Kristeller, J. & Merriman, P. (2008). Mindfulness, spirituality, and health-related symptoms. *Journal of Psychosomatic Research*, 64, 393–403.

Cebolla, A., García-Palacios, A., Soler, J., Guillen, V., Baños, R. & Botella, C. (2012). Psychometric properties of the Spanish validation of the five facets of mindfulness questionnaire (FFMQ). *European Journal of Psychiatry*, 26, 118–126.

Chadwick, P., Hember, M., Symes, J., Peters, E., Kuipers, E. & Dagnan, D. (2008). Responding mindfully to unpleasant thoughts and images: Reliability and validity of the Southampton mindfulness questionnaire (SMQ). *British Journal of Clinical Psychology*, 47, 451–455.

Cheung, G. W. & Rensvold, R. B. (2002). Evaluating goodness-of-fit indexes for testing measurement invariance. *Structural Equation Modelling: A Multidisciplinary Journal*, 9, 233–255.

Clark-Carter, D. (2010). *Quantitative psychological research: The complete student's companion* (3rd edn). Hove, UK: Psychology Press.

Coffey, K. A., Hartman, M. & Fredrickson, B. L. (2010). Deconstructing mindfulness and constructing mental health: Understanding mindfulness and its mechanisms of action. *Mindfulness*, 1, 235–253.

Dundas, I., Vøllestad, J., Binder, P. E. & Sivertsen, B. (2013). The five factor mindfulness questionnaire in Norway. *Scandinavian Journal of Psychology*, 54, 250–260.

Eriksson, M. & Lindström, B. (2005). Validity of Antonovsky's sense of coherence scale: A systematic review. *Journal Epidemiology Community Health*, 59, 460–466.

Falkenström, F. (2010). Studying mindfulness in experienced meditators: A quasi-experimental approach. *Personality and Individual Differences*, 48, 305–310.

Feldman, G., Hayes, A., Kumar, S., Greeson, J. & Laurenceau, J.-P. (2007). Mindfulness and emotion regulation: The development and initial validation of the Cognitive and Affective Mindfulness Scale-Revised (CMS-R). *Journal of Psychopathology and Behavioral Assessment*, 29, 177–190.

Fresco, D. M., Moore, M. T., van Dulmen, M. H. M., Segal, Z. V., Ma, S. H., Teasdale, J. D. & Williams, J. M. G. (2007). Initial psychometric properties of the experiences questionnaire: Validation of a self-report measure of decentering. *Behavior Therapy*, 38, 234–246.

Gu, J., Strauss, C., Bond, R. & Cavanagh, K. (2015). How do mindfulness-based cognitive therapy and mindfulness-based stress reduction improve mental health and wellbeing? A systematic review and meta-analysis of meditation studies. *Clinical Psychology Review*, 37, 1–12.

Haigh, E. A. P., Moore, M. T., Kashdan, T. B. & Fresco, D. M. (2011). Examination of the factor structure and concurrent validity of the Langer Mindfulness/Mindlessness Scale. *Assessment*, 18, 11–26.

Hayes, S. (2002). Buddhism and acceptance and commitment therapy. *Cognitive and Behavioral Practice*, 9, 58–66.

Hermann, C. (1997). International experiences with the hospital anxiety and depression scale – a review of validation data and clinical results. *Journal of Psychosomatic Research*, 42, 17–41.

Jensen, C. G., Krogh, S. C., Westphael, G. & Hjordt, L. V. (2019). Mindfulness is positively related to socioeconomic job status and income and independently predicts mental distress in a long-term perspective: Danish validation studies of the five-factor mindfulness questionnaire. *Psychological Assessment*, 1, e1–e20.

Kabat-Zinn, J. (2013). *Full catastrophe living. Using the wisdom of your body and mind to face stress, pain and illness* (2nd rev. edn). New York: Bentham Random House.

Kiken, L. G., Garland, E. L., Bluth, K., Palsson, O. S. & Gaylord, S. A. (2015). From a state to a trait: Trajectories of state mindfulness in meditation during intervention predict changes in trait mindfulness. *Personality and Individual Differences*, 81, 41–46.

Kline, RB. (2005). *Principles and practice of structural equation modelling* (2nd edn). New York: Guilford Press.

Lau, M. A., Bishop, S. R., Segal, Z. V., Buis, T., Anderson, N. D., Carlson, L. *et al.* (2006). The Toronto mindfulness scale: Development and validation. *Journal of Clinical Psychology*, 62, 1445–1467.

Lilja, J. L., Frodi-Lundgren, A., Johansson Hanse, J., Josefsson, T., Lundh, L.-G., Sköld, C. *et al.* (2011). Five facet mindfulness questionnaire – reliability and factor structure: A swedish version. *Cognitive Behaviour Therapy*, 40, 291–303.

Lilja, J. L., Lundh, L.-G., Josefsson, T. & Falkenström, F. (2013). Observing as an essential facet of mindfulness: A comparison of FFMQ patterns in meditating and non-meditating individuals. *Mindfulness*, 4, 203–212.

Lilja, J. L., Zelleroth, C., Axberg, U. & Norlander, T. (2016). Mindfulness-based cognitive therapy is effective as relapse prevention for patients with recurrent depression in Scandinavian primary health care. *Scandinavian Journal of Psychology*, 57, 464–472. https://doi.org/10.1111/sjop.12302.

Lisspers, J., Nygren, A. & Soederman, E. (1997). Hospital Anxiety and Depression Scale (HAD): Some psychometric data for a Swedish sample. *Acta Psychiatrica Scandinavica*, 96, 281–286.

Muthén, B. O. & Asparouhov, T. (2012). Bayesian structural equation modelling: A more flexible representation of substantive theory. *Psychological Methods*, 17, 313–335.

Muthén, L. K. & Muthén, B. O. (1998–2017). *Mplus user's guide*. Los Angeles, CA: Muthén & Muthén.

Olsson, M., Gassne, J. & Hansson, K. (2009). Do different scales measure the same construct? Three sense of coherence scales. *Journal of Epidemiology and Community Health*, 63, 166–167.

Park, T., Reilly-Spong, M. & Gross, C. R. (2013). Mindfulness: A systematic review of instruments to measure an emergent patient reported outcome (PRO). *Quality of Life Research*, 22, 2639–2659.

Pearson, M. R., Lawless, A. K., Brown, D. B. & Bravo, A. J. (2015). Mindfulness and emotional outcomes: Identifying subgroups of college students using latent profile analysis. *Personality and Individual Differences*, 76, 33–38.

Pett, M. A., Lackey, N. R. & Sullivan, J. J. (2003). *Making sense of factor analysis: The use of factor analysis for instrument development in health care research*. London: SAGE Publications.

Segal, Z. V., Williams, J. M. G. & Teasdale, J. D. (2002). *Mindfulness-Based Cognitive Therapy for depression: A new approach to preventing relapse*. New York: Guilford Press.

Shahar, B., Britton, W. B., Sbarra, D. A., Figueredo, A. J. & Bootzin, R. R. (2010). Mechanisms of change in mindfulness-based cognitive therapy for depression: Preliminary evidence from a randomized controlled trial. *International Journal of Cognitive Therapy*, *3*, 402–418.

Shapiro, S. L., Oman, D., Thoresen, C. E., Plante, T. G. & Flinders, T. (2008). Cultivating mindfulness: Effects on well-being. *Journal of Clinical Psychology*, *64*, 840–862.

Van Dam, N. T., Hobkirk, A. L., Danoff-Burg, S. & Earleywine, M. (2012). Mind your words: Positive and negative items create method effects on the five facet mindfulness questionnaire. *Assessment*, *19*, 198–204.

Veehof, M. M., ten Klooster, P. M., Taal, E., Westerhof, G. J. & Bohlmeijer, E. T. (2011). Psychometric properties of the Dutch Five Facet Mindfulness Questionnaire (FFMQ) in patients with fibromyalgia. *Clinical Rheumatology*, *30*(8), 1045–1054. https://doi.org/10.1007/s10067-011-1690-9.

Zigmond, A. S. & Snaith, R. P. (1983). The Hospital Anxiety and Depression Scale. *Acta Psychiatrica Scandinavica*, *67*, 361–370.