EVALUATING THE PSYCHOMETRIC PROPERTIES OF THE FIVE-FACET MINDFULNESS QUESTIONNAIRE: AN ITEM RESPONSE THEORY ANALYSIS

by

Angela S. Bowman

A Thesis Submitted in Partial Fulfillment
of the Requirements for the Degree of
Master of Arts in Quantitative Psychology

Middle Tennessee State University
May 2014

Thesis Committee

Dr. Jwa K. Kim

Dr. Dana K. Fuller

Dr. Catherine E. Crooks

This research is dedicated to my son, W. Jeremy Bowman, and to my grandmother, Sharon Floyd; you are the source of my perseverance and desire to always push myself to "do better."

ACKNOWLEDGEMENTS

First and foremost, I must praise God for the many blessings He has bestowed on me, including placing the people in my life without whom this research would not have been possible. I would like to express my sincerest appreciation to my thesis advisor Dr. Jwa Kim for his support, patience, depth of knowledge of psychometrics, and for his jokes whenever things were getting too serious. His guidance provided direction and aided in a deeper understanding of psychometric analysis. He also taught me that no matter how many jokes you tell, the problem will still be on the board. I would also like to thank Dr. Dana Fuller who has been a constant source of inspiration due to her commitment to excellence, depth of knowledge of quantitative analysis, and strong sense of integrity. Finally, I would to thank Dr. Catherine Crooks whose enthusiasm and support kept me moving forward and fueled my motivation to dig deeper.

I would like to thank my colleague and friend Michelle Sterlingshires who has been a constant source of support, a wealth of knowledge, and always willing listen and offer suggestions. I must also express heartfelt gratitude to my closest friend Amanda Clevinger. Amanda has been my rock, my biggest cheerleader, and my guiding light when I was convinced there was no hope of surviving graduate school. Finally, I would like to express my deepest love and gratitude to my husband Mark who gave up everything in order allow me to pursue my dreams, to my parents and younger brothers who have been a constant source of encouragement, and to my dear friend Mary Mulligan without whom I would have never been able to move to Tennessee.

ABSTRACT

Mindfulness has been extensively studied and a plethora of research delineates the relationship between mindfulness well-being. The Five-Facet Mindfulness Questionnaire (FFMQ) measures the dominant facets of mindfulness. The FFMQ has been shown to have solid classical test properties, but often the factor structure does not hold up under factor analysis thus an item analysis was conducted. Data were analyzed from 644 individuals. Reliability of each facet of the FFMQ (observing, describing, acting with awareness, non-judging of inner experiences, and non-reactivity to inner experiences) were found to have mediocre to good classical test properties. Under Samejima's graded response model, the pattern of results revealed poor discrimination and little information provided at all levels of the trait for each item of the observing and non-reactivity to inner experiences scales. However, the remaining scales appeared to discriminate well and provided adequate information at a range the trait. Limitations and future directions are discussed.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

CHAPTER I

Introduction

Mindfulness has been defined as a multi-faceted construct characterized by an attentional shift that results in the observation of moment to moment experiences without interpretation, elaboration, or analysis of the experience, a non-judgmental awareness of the here and now (Kabat-Zinn, 2000). More specifically, mindfulness is "paying attention in a particular way: on purpose, in the present moment, and nonjudgmentally" (Kabat-Zinn, 1994, p. 4). Meditation has been at the core of Buddhist tradition for thousands of years, and recently has become a popular topic of Western psychological research as it becomes evident that mindfulness based interventions produce clinically significant changes in depression and anxiety (Baer, 2011). Meditation is merely the scaffolding upon which the skill of mindfulness is built. The primary focus of Western psychology has been on integrating mindfulness meditation into both prevention and intervention treatments designed to reduce stress levels, ease symptoms of psychological distress, and improve the individual's sense of psychological well-being (Kabat-Zinn, 1990).

Mindfulness and Well-being

A plethora of studies have examined the relationship between mindfulness, perceived stress, and psychological well-being. According to Leary (2004), the majority of mental health theories assume that having a factual perception of reality is a classic sign of psychological well-being. Brown and Ryan (2003) also argue that mindfulness facilitates well-being directly by adding a sense of clarity to experiences by encouraging

"moment to moment non-judgmental contact with life" (p. 224). In a 2008 study by Shapiro, Oman, Thoresen, Plante, and Flinders, mindfulness was shown to have both direct and mediating effects on the reduction of negative rumination and perceived stress, which led to a significant increase in subjective well-being.

Mindfulness is thought to facilitate a change in the regulation of attention, cognition and emotion by discouraging habitual automatic thought patterns such as rumination and obsession (Teasdale et al., 2002). Higher scores on mindfulness scales, such as the Mindful Attention Awareness Scale (MAAS), have been associated with lower levels of depression, anxiety, and stress symptoms, as well as being associated with increased subjective and eudemonic well-being (Brown & Ryan, 2003). Participation in Mindfulness-Based Stress Reduction (MBSR) programs has proven to be effective in reducing self-reported distress and mood disturbances while increasing trait mindfulness (Baer, Smith, Hopkins, Krietemeyer, & Toney, 2006). Additionally, in those who have experienced more than three episodes of depression, MBSR participation has been shown to reduce that rate at which relapse occurs (Teasdale et al., 2000), as well as reduce suicidal ideation and facilitate social adjustment of those diagnosed with borderline personality disorder (Turner, 2000). Brown, Ryan, and Creswell (2007) provided an extensive review of additional studies outlining the correlation between self-reported mindfulness and various psychological constructs including subjective well-being and psychopathology.

Measuring Mindfulness

Given the positive effects of mindfulness on well-being, several self-report measures to quantify the state (present moment) and trait (stable and enduring) qualities of mindfulness have emerged. Many of these scales have been shown to significantly correlate with one another (Baer, Smith, Hopkins, Krietemeyer, & Toney, 2006). The Freiburg Mindfulness Inventory (FMI) is a measure derived from the Buddhist roots of mindfulness. It consists of 30 items designed for use with experienced meditators, and measures one's openness to negative experience and nonjudgmental present-moment observation (Buchheld, Grossman, & Walach, 2001). The FMI was designed to measure mindfulness as a multifaceted construct with four dominant factors, but the four-factor pattern has proven to be unstable in validation studies (Walach, Buchheld, Buttenmüller, Kleinknecht, & Schmidt, 2006). However, Buchheld et al. (2001) reported excellent internal consistency with Cronbach's alpha ranging from .93 to .94 in pre and post mindfulness training assessment.

The Mindful Attention Awareness Scale (MAAS) contains 15 items designed to measure the single latent construct of one's natural tendency to be aware of, or attend to, present-moment experience (Brown & Ryan, 2003). Whereas all items of the FMI are positively worded, the MAAS contains negatively worded items. Therefore, the MAAS could be said to measures one's tendency to act with automaticity, or lack of attention to detail. However, the MAAS does offer a measure of mindfulness intended for use in non-meditating populations. Brown and Ryan (2003) reported good internal consistency with Cronbach's coefficient alpha of .87.

The Cognitive and Affective Mindfulness Scale (CAMS) is a 12-item measure consisting of items intended to assess present centered focus, attention, awareness, and nonjudgmental acceptance of emotions and thoughts in day-to-day life (Hayes & Feldman, 2004). The CAMS differs from other measures of mindfulness in that items attempt to capture one's ability or inclination to be mindful in lieu of daily level of mindfulness (Bergomi, Tschacher, & Kupper, 2013). Additionally, the CAMS was designed to be relevant in the treatment of depression (Hayes & Feldman, 2004). The authors found the CAMS to have acceptable internal consistency with coefficient alpha ranging from .74 to .80.

The Southampton Mindfulness Questionnaire (SMQ) consists of 16 items and is designed to measure if one has a mindful approach for dealing with distressing thoughts and/or images (Chadwick et al., 2008). The authors reported good internal consistency with coefficient alpha of .89. As with the FMI, the SMQ was designed to measure mindfulness as a multifaceted construct, but has proven to measure a single trait. The SMQ focuses on distressing internal events, which allows it to be a useful tool in assessing the relationship between mindfulness and aspects of mental health. However, the focus on distressing events limits the usefulness of the scale in non-clinical populations due to a lack of items measuring positive or neutral experiences.

The final scale to be discussed here is the Kentucky Inventory of Mindfulness Skills (KIMS) which is a 39-item measure designed to assess four facets associated with mindfulness in daily life: observing, describing, acting with awareness, and accepting without judgment (Baer, Smith, & Allen, 2004). The KIMS was developed in the

framework of Dialectical Behavioral Therapy (DBT; Linehan, 1993), and is unique in that it incorporates a measure of *describing*. Mindfulness training is one aspect of DBT where the practice of observing one's environment and using nonjudgmental labels to describe what has been observed is encouraged. Internal consistency for the four sub-scales of the KIMS range from a reported Cronbach's alpha of .76 to .91. Example items from each of the aforementioned mindfulness scales can be seen in Table A1, in the Appendix.

Development of the FFMQ

Baer, Smith, Hopkins, Krietemeyer, and Toney (2006) combined the items of the above-mentioned mindfulness scales (CAMS, SMQ, FMI, MAAS, and KIMS) in order to conduct an exploratory factor. Results of principal axis factoring using oblique rotation produced 26 factors with eigenvalues greater than 1.0 that accounted for 63% of the total variance. Because using eigenvalues greater than 1.0 can result in an overestimate of the number of factors (Floyd & Widamon, 1995), Baer et al. (2006) opted to rely on the scree plot for determining the number of factors to retain, which revealed a five-factor structure that appeared to capture the underlying characteristics of mindfulness. A second EFA forcing a five-factor structure was conducted and the five-factors retained accounted for 33% of the total variance. Baer et al. (2006) identified these five factors as: *Observing* (attending to internal or external stimuli), *Describing* (mentally labeling stimuli), *Acting with Awareness* (consciously attending to actions without behaving with automaticity), *Non-Judging of Inner Experiences* (reframing from evaluation of thoughts, sensations,

and feelings) and *Non-Reactivity to Inner Experiences* (allowing thoughts and feelings to arise and dissipate without rumination).

Thirty-nine items with no cross-loadings and minimum factor loading of .40 were retained and combined to create the Five-Facet Mindfulness Questionnaire (FFMQ). To create the FFMQ, Baer et al. (2006) retained eight items for each of the four factors related to the KIMS and seven items related to the one factor from the FMI and SMQ. Four of the five factors: observing, describing, acting with awareness, and non-judging of inner experiences, were similar to the factor structure of the KIMS (Baer et al. 2004), while the fifth factor, non-reactivity to inner experiences, consisted of items found in the FMI and SMQ. The internal consistency of the five subscales was measured and revealed the following Cronbach alpha values: Non-Reactivity to Inner Experiences = .75, Observing = .83, Acting with Awareness = .87, Non-Judging of Inner Experiences = .87, and Describing = .91. Therefore, all five subscales were determined to be of adequate to good consistency.

After arriving at a five-facet structure of mindfulness, Baer et al. (2006) conducted a confirmatory factor analysis (CFA), with an independent sample, to investigate and replicate the proposed structure. Item parceling, based on methods suggested by Little, Cunningham, Shahar, and Widamon (2002), was used and two parceling procedures were tested. In the first analysis, items were assigned to parcels randomly. Items were then reassigned to parcels based on factor loadings. In both parceling methods, each factor contained three parcels. Details of which items were

assigned to which parcels were not provided, nor were detailed results of the two parceling methods.

Baer et al. (2006) tested several models, including a one-factor model, a five-factor correlated model, and a hierarchical model. The one-factor model produced abysmal fit indices, while the five-factor correlated model produced a strong fit. The hierarchical structure, where the five factors were components of a mindfulness construct, fit the data moderately well, but provided a reduced fit over the five-factor correlated model. Details of model fit for each of the three models tested can be seen in Table A2 in the Appendix.

Validation of the FFMQ

Baer et al. (2008) conducted a follow up study to assess the validity of the five-factor structure. Data were gathered from a meditating sample and used in comparison to the 2006 non-meditating sample. As with the 2006 study, item parceling was used to improve model fit by reducing the number of estimated parameters. The results of a CFA conducted on the meditating sample data revealed that both a five-factor correlated structure and a hierarchical structure fit the data well with very little difference between the two models. Details of model fit can be seen in Table A2, while the final version of the FFMQ can be seen in Table A3.

In the creation (Baer et al., 2006) and validation (Baer et al., 2008) of the FFMQ, item parceling was used to improve model fit. However, no information was provided for model fit without parceling the data. According to Bandalos (2002), item parceling

impacts model fit by reducing the sample covariance matrices and thus reducing error components from unmodeled associations within those matrices. The result is improved model fit for a potentially mispecified model. The author suggested that item parceling is acceptable in various circumstances (e.g., when the assumption of local independence is violated; with dichotomous aptitude and achievement measures). However, she suggested that item parceling should be avoided in attitude and personality research where improved model fit should take a backseat to understanding the nature of relationships among variables (Bandalos, 2002).

Given the pitfalls of item parceling, Christopher, Neurser, Nivhaiel, and Baitmangalkar (2012) analyzed the properties of the FFMQ without the use of item parceling using a demographically similar combined sample of meditators and non-meditators. They conducted a confirmatory factor analysis without using item parcels and found that the data from the FFMQ best fit a five-factor hierarchical model with excellent fit indices and strong factor loadings; details appear in Table A2. In addition, they provided support for incremental validity of the FFMQ. Scores on the FFMQ were positively correlated with The Satisfaction with Life Scale (r = .52) and negatively correlated with The Center for Epidemiologic Studies Depression Scale (r = -.58). Finally, after controlling for age and education, Describing, Acting with Awareness, Non-Reactivity to Inner Experiences, and Non-Judging of Inner Experiences were positively correlated with satisfaction with life, while Acting with Awareness, Non-Judging of Inner Experiences, and Non-Reactivity to Inner Experiences were negatively correlated with depression. Although the previously mentioned studies provide support

for the FFMQ as a measure of mindfulness in various populations, studies examining the properties of the FFMQ in non-meditating clinical populations often result in reduced model fit.

Recently, Bohlmeijer, ten Klooster, Fledderus, Veehof, and Baer (2011) sought to validate the FFMQ in a non-meditating Dutch sample of individual's suffering from mild to moderate depression. Although, the subscales produced similar alpha levels and intercorrelations as studies in non-clinical samples, the model fit indices were marginally lower than those found by Baer et al. (2006; 2008) and Christopher et al. (2012). Similarly, Hou, Wong, Lo, Mak, and Ma (2013) analyzed a Chinese version of the FFMQ in a non-meditating sample and produced results similar to those of Bohlmeijer et al. (2011). Further examples include Veehof, ten Klooster, Taal, Westerhof, and Bohlmeijer (2011) and Fernandex, Wood, Stein and Rossi (2010). The intercorrelations between facets, coefficient alpha levels for subscales, and model fit indices can be seen in Table A2.

Principles of Item Response Theory

In prior studies, psychometric properties of the FFMQ have been reported using classical test theory (CTT). In CTT, an individual's score is a function of their true score plus error in measurement (Lord & Novick, 1969). Although CTT is a valuable tool in scale development, it has several problems and limitations that may explain the variation in results of prior mindfulness studies. Hambleton, Swaminathan, and Rogers (1991) outline many of the limitations of CTT. The estimates of item parameters are *group-dependent*, meaning that results are not generalizable to a population that differs from the

sample measured. While the estimates of an individual's true scores are *test-dependent*, meaning that even when the same scale is administered, "scores are unequally precise" (p. 4) at differing latent trait levels (Hambleton, Swaminathan, & Rogers, 1991). Thus in CTT separating Yet another limitation of CTT discussed by Hambleton et al. (1991) is that it is test driven instead of item driven. Therefore, no information is provided as to how individuals with different levels of mindfulness perform on each item on the FFMQ. Finally, CTT presumes the existence of parallel test forms, which in reality is an impossible assumption to meet.

In item response theory (IRT), the true score of CTT is replaced with a person parameter, while item and test indices are replaced with item parameters that are group invariant (Hambleton et al., 1991). Therefore, one of the major benefits of IRT is that item parameter estimates need not to be replicated for varying populations. In contrast to CTT, IRT links the probability of a response on an item to the individual's level of the latent trait being measured and this relationship is described by the item characteristic curve (Hambleton et al., 1991). Therefore, regardless of the population from which an individual originates, the probability of a specific response to an item on the FFMQ is, in theory, a direct function of their level of mindfulness. An example of the item characteristic curve, where the level of the latent trait (ability) is on the X-axis and the probability of correct response is on the Y-axis, can be seen in Figure A1.

Another advantage of IRT is a reduction in the number of assumptions that must be met, namely unidimensionality and local independence (Lord & Novick, 1969). Unidimensionality is met when there is a single dominant latent trait being measured, and

local independence is met when the responses to any two items are statistically independent when the latent trait is accounted for. Unidimensionality is a vital assumption of IRT, but it is often a difficult assumption to meet due to influences from cognitive factors, testing effects, or personality (Hambleton et al., 1991). However, IRT models have been shown to be robust to violations of unidimensionality when there is a relatively small secondary factor or if two factors are highly correlated (Dragsow & Parsons, 1983; Reckase, 1979; Reckase & McKinley, 1991).

Logistic IRT Models

According to Hambleton et al. (1991), the most commonly used IRT models are the one, two, and three parameter logistic models. Logistic models are useful for estimating item parameters for both dichotomous and polytomous scored items. In the one-parameter logistic model the *b-parameter* is the estimate of interest. The *b-parameter* is typically referred to as item difficulty and as such is an estimate of the level of latent trait present at the .50 probability of a correct response. Therefore, the one-parameter logistic model (1PL) assumes that the probability of answering an item correctly is a function of the level of the latent trait and item difficulty. According to Birnbaum (1968), the form of the item characteristic curve for the 1, 2, and 3PL models are given as follows:

1. For the one-parameter logistic model:

$$P_i(\theta) = \frac{e^{(\theta - b_i)}}{1 + e^{(\theta - b_i)}} \qquad i = 1, 2, \ldots, n \qquad (1)$$

where:

$P_i(\theta)$ is the probability that a randomly chosen examinee with ability θ answers item *i* correctly,
b_i is the item *i* difficulty parameter,
n is the number of items in the test,
e is the constant 2.718, and,
$P_i(\theta)$ is an S-shaped curve with values between 0 and 1 over the ability scale.

2. For the two-parameter logistic model:

$$P_i(\theta) = \frac{e^{D_{a_i}(\theta - b_i)}}{1 + e^{D_{a_i}(\theta - b_i)}} \qquad i = 1, 2, \ldots, n \qquad (2)$$

where:

D is a scaling factor used to make the logistic function as close as possible to the normal ogive function, where D = 1.7, and
a is the item *i* discrimination parameter.

3. For the three-parameter logistic model:

$$P_i(\theta) = c_i + (1 - c_i)\frac{e^{D_{a_i}(\theta - b_i)}}{1 + e^{D_{a_i}(\theta - b_i)}} \qquad i = 1, 2, \ldots, n \qquad (3)$$

where:

c_i is a *pseudo-chance-level* (guessing) parameter.

When interpreting item difficulty for the logistic models, it is important to keep in mind that some point of reference is required. What is "difficult" for one population may not be so for another population. Therefore, when interpreting item difficulty, it must be in relation to the location of the .50 probability of correct response on the ability scale (Baker, 2001). Interpreting item discrimination is a bit more straightforward. The range of values for interpreting item discrimination for the logistic model can be seen in Table A4.

Samejima's Graded Response Model for Polytomous Items

Samejima (1969) outlined a model for use with polytomous items with ordered categorical responses (e.g., Likert). According to Embretson and Reise (2000), Samejima's graded response model (SGRM) is an indirect model in that a two-step process is utilized to denote the conditional probability for an examinee to have a response beyond a particular category. SGRM is a generalization of the two-parameter dichotomous logistic model in that it treats two adjacent response categories (m) as dichotomous categories (m and m - 1). Each item has one common slope parameter (a_i), and m - 1 category thresholds (β_{ij}; Embretson & Reise, 2000). The goal of SGRM is the find the point on the latent trait continuum that corresponds to each category threshold. According to Embretson and Reise (2000), "the probability of an examinee's raw score (x) falling in or above a given category threshold ($j = 1, \ldots, m_i$), conditional on trait level (θ)" (p. 96) is given by:

$$P_{ix}^{*}(\theta) = \frac{e^{a_{ij}(\theta-\beta_{ij})}}{1+e^{a_{ij}(\theta-\beta_{ij})}} \qquad x = j = 1, 2, \ldots, m_i \qquad (4)$$

where:

a_{ij} is a common slope parameter, and
β_{ij} is the category threshold parameter.

The curves created by the $P_{ix}^{*}(\theta)$ function are referred to as operating characteristic curves (OCC). One curve must be calculated for each between category threshold location (β_{ij}). Thus, for a measure with a 5-point Likert scale, four β_{ij} and four OCC's must be computed. Each β_{ij} denotes where on the latent trait continuum (θ) an examinee must fall in order to reach the 50% probability of scoring above response category *j*. In SGRM each item has a series of $m = K - 1$ dichotomies, and parameter estimates are generated for each dichotomy using the two-parameter logistic model with the slope parameter held constant for each curve (Embretson & Reise, 2000).

One the OCC's have been estimated the category response functions (CRF) can be computed. The CRF displays one curve for each response category with the probability of a response on the Y-axis and the latent trait on the X-axis. The probability of the lowest response is the probability of not passing the first threshold (monotonically decreasing), while the probability of giving the next highest response is the probability of having crossed the first response threshold. This pattern follows for each increasing category to the highest category where the probability of a response is monotonically increasing (Samejima, 1969). The CRF is computed by subtraction and given by:

$$P_{ix}(\theta) = P^{*}_{ix}(\theta) - P^{*}_{i(x+1)}(\theta) \quad (5.1)$$

where:

$P^{*}_{i0} = 1.0$ is the probability of responding in the lowest category monotonically decreasing, and

$P^{*}_{i5} = 0.0$ is the probability of responding above the highest category monotonically increasing.

Thus:

$$P_{i0}(\theta) = 1.0 - P^{*}_{i1}(\theta) \quad (5.2)$$

$$P_{i1}(\theta) = P^{*}_{i1}(\theta) - P^{*}_{i2}(\theta) \quad (5.3)$$

$$P_{i2}(\theta) = P^{*}_{i2}(\theta) - P^{*}_{i3}(\theta) \quad (5.4)$$

$$P_{i3}(\theta) = P^{*}_{i3}(\theta) - P^{*}_{i4}(\theta) \quad (5.5)$$

$$P_{i4}(\theta) = P^{*}_{i4}(\theta) - P^{*}_{i5}(\theta) \quad (5.6)$$

In Samejima's graded response model, the location and shape of the OCC and CRF are determined by the item parameter estimates. The locations of each OCC, and of the peak of the curves for the CRF of the middle categories are determined by the values of β_{ij}. Additionally, how narrow or peaked the CRF's will be and how steep the OCC will be are determined by a_{ij}. In general, a CRF is expected to peak in the center of two category thresholds with higher slope parameters producing more narrow and peaked CRF's that signify category responses that discriminate among levels of the latent trait (Embretson & Reise, 2000).

Purpose of the Current Study

Although researchers examining the FFMQ in the framework of CTT have agreed that the scale has sound psychometric properties, the proposed model of mindfulness from which the scale is derived varies in how well the data fit the model depending on

the population from which the sample is drawn. It is unknown if these differences are due to model misspecification or potentially problematic items within facets of the FFMQ. To date, an item response analysis has not been conducted. In order to understand the underlying characteristics of mindfulness, it is important to understand how each item of the scale functions, independent of scale and sample characteristics. Therefore, the goal of the current study is to examine each item of the FFMQ in the framework of item response theory (IRT) utilizing Samejima's Graded Response Model (SGRM).

CHAPTER II

Methods

Participants

Upon approval of the institutional review board, participants were recruited from the undergraduate research pool at a Southeastern University and received one research credit for their participation. A total of 735 individuals completed the informed consent and survey material. Prior to any analyses, the data were examined for patterns of missingness using Little's MCAR Test (Little, 1998). According to Little (1998), a non-significant χ^2 value indicates that data are missing completely at random (MCAR) and therefore missingness can be ignored. Little's MCAR Test indicated that the data were missing completely at random, χ^2 (1472, n = 735) = 1540.72, p = .104 and thus missing data were handled through the use of multiple imputation (Rubin, 1987).

Items included as a validity check were examined for correct responses. Twenty-two participants were removed for failing to respond to a minimum of 75% of items overall and 90% of items on the FFMQ. An additional 69 participants were removed for failing to pass the validity questions, resulting in a final sample of 644 respondents. Sixty-eight percent of respondents were female (n = 438), and the average age was 20 years (SD = 4). In regards to ethnicity, 59.5% (n = 383) indicated they were White, 27.2% (n = 175) were Black, 3.9% (n = 25) were Hispanic, 3.7 (n = 24) were Asian, 0.8% (n = 5) were Native American, and 4.6% (n = 30) selected "other". When asked about experience with mediation, 20.7% (n = 133) indicated that they were currently meditating.

Of those who indicated that they were currently meditating, 77.5% (*n* = 103) indicated that they meditated three days or less per week, while 12.0% (*n* = 16) indicated having had some form of formal meditation training.

Materials

Demographic Information. Participants completed a demographics form requesting their age, gender, ethnicity, and experience with meditation. Demographic questions can be seen in Appendix B.

Five Facet Mindfulness Questionnaire. (FFMQ; Baer et al. 2006) is a 39-item likert scale designed to assess the five factors associated with the tendency to be mindful in everyday activities (i.e., Acting with Awareness, Describing, Non-Judging of Inner Experiences, Non-Reactivity to Inner Experiences, and Observing). Participants were asked to rate if each item is generally true for them on a five-point scale from "Never or Very Rarely True" to "Very Often or Always True." Internal consistencies of the FFMQ subscales based on Cronbach's Alpha were as follows: Observing = .83, Describing = .91, Acting with Awareness = .87, Non-Judging of Inner Experiences = .87, and Non-Reactivity to Inner Experiences = .75, with full scale internal consistency of .86 in a non-meditating sample and .95 in a meditating sample. The items of the FFMQ can be seen in Table A3.

Depression Anxiety Stress Scale. Depression Anxiety Stress Scale (DASS-42; Lovibond & Lovibond, 1993) is a 42-item likert scale designed to assess negative emotional states associated with depression, anxiety, and stress over the previous two

weeks. Each item was scored on a four-point likert scale, from "Did not apply to me at all" to "Applied to me very much, or most of the time." The scores for depression, anxiety, and stress were calculated by summing the 14-items relevant to the characteristic being measured. According to Lovibond and Lovibond (1993), internal consistencies based on Cronbach's Alpha for the DASS were as follows: Depression .91, Anxiety .84, and Stress .90. Items of the DASS-42 can be seen in Appendix B.

Procedure

All data were collected online using the online survey utility, Qualtrics (Qualtrics, Provo, UT). Upon obtaining informed consent, each participant was asked to complete demographic information, the FFMQ and the DASS, followed by a short debriefing. Four items asking the participant to select a particular response were included as a validity check.

Analysis

Data cleaning. The data were analyzed using a combination of IBM SPSS v20 (IBM Corp., 2011) and Xcalibre v4.2 (Assessment Systems Corporation, 1996). Items included as a validity check were examined for correct responses, and participants who failed to produce correct responses were not further analyzed. Additional participation criteria included responding to a minimum of 75% of items overall, and 90% of the items on the FFMQ.

Scale scores were tabulated by summing the total of all FFMQ items. Additionally, subscale sores for the FFMQ and DASS were created by summing the

scores of associated items. Based on prior research and theory, the Observing scale consisted of items 1, 6, 11, 15, 20, 26, 31, and 36. The Describing scale consisted of items 2, 7, 12, 16, 22, 27, 32, and 37. The Acting with Awareness scale consisted of items 5, 8, 13, 18, 23, 28, 34, and 38. The Non-Judging of Inner Experiences scale consisted of items 3, 10, 14, 17, 25, 30, 35, and 39. The Non-Reactivity to Inner Experiences scale consisted of items 4, 9, 19, 21, 24, 29, and 33. Additionally, the incremental validity of the FFMQ was assessed by examining the correlations between the scale mindfulness score and scores on the depression and anxiety subscales of the DASS, in addition to examining the correlation between age and mindfulness. A pattern of correlations similar to those found in prior studies was expected.

Classical Test Theory (CTT). Coefficient alpha for each of the five subscales of the FFMQ, as well as the full scale, was calculated to determine the reliability (Cronbach, 1951). Values similar to those reported in prior studies were expected, with values greater than .70 considered acceptable for research purposes (Nunnaly, 1978). Corrected item to subscale correlations were examined, as well as expected change in alpha if the item was to be removed.

Exploratory Factor Analysis (EFA). Prior to conducting an item response analysis the unidimensionality of each subscale was assessed. An exploratory factor analysis (EFA) with maximum likelihood (ML) estimation (Joreskog, 1969) was conducted, with the expectation of unidimensional structure in each subscale and a five-facet structure in the full FFMQ. Bartlett's Test of Sphericity and the Kaiser-Meyer-Olking measure of sampling adequacy were assessed to determine the appropriateness of

factor analysis. Bartlett's Test of Sphericity tests the null hypothesis that the intercorrelation matrix (S) is derived from a population where the intercorrelation matrix (Σ) is an identity matrix (Bartlett, 1954). Therefore, a significant chi square value was expected and indicative that the non-zero correlations of the S-matrix are not due to sampling error and factor analysis is appropriate. Given that statistics based on chi square are sensitive to sample size, the Keiser-Meyer-Olken Measure of Sampling Adequacy (KMO) was also assessed. The KMO assesses the sum of the squared partial correlations in relation to the sum of the squared correlations, with values closer to 1.0 indicative of small partial correlations (Kaiser, 1974). The value of the KMO can range from 0 to 1 with values above .60 indicating factorability (Tabachnick & Fidell, 2013).

In order to determine dimensionality a combination of the scree plot, eigenvalues, and factor loadings was examined. According to Kaiser (1960), factors with eigenvalues greater than 1.0 that explain greater than 5-10% of the total variance should be retained. However, a cut-off of 1.0 can result in retaining arbitrary factors that are the result of random noise, while a cut-off of 1.40 has proven to be more stable, and thus a cut-off of 1.40 was used in this study (Smith & Miao, 1994). The scree plot of eigenvalues was examined to determine the point of inflection. According to Cattell (1966), only those values that fall above the point of inflections should be retained. Finally, the values of the factor loadings for each item was considered. Loadings greater than .71 are considered excellent, values around .63 are considered very good, values around .55 are considered good, values around .45 are considered fair, and values around .32 or below are considered poor (Comrey & Lee, 1992). Therefore, the number of non-trivial factors

was determined through a combination of eigenvalues greater than 1.40, the inflection point of the scree plot, and factor loadings greater than .45.

Item Response Theory (IRT). Each item of the FFMQ was analyzed using the principles of item response theory (IRT), using a Samejima's graded response model (SGRM; Samejima, 1969). Estimates were generated using expected a posteriori (EAP) estimation (Uebersax, 1993). Chernyshenko, Stark, Chan, Dragsow, and Williams (2001) point out that measures of attitude and personality do not perform in a similar fashion as measures of cognitive ability under SGRM. Through empirical and simulation research, it was determined that model/data and model/item fit were best measured using the chi-square/*df* ratio rather than relying on chi-square values alone. The overall model fit (chi-square/*df* ratio) was examined to determine the appropriateness of the SGRM model and was expected to be below 3.0 per the recommendations of Chernyshenko et al. (2001). The test information function (TIF) was examined and due to the nature of the scale, it was expected to indicate that unique information was provided at various level of the latent trait. Similarly, item information functions (IIF) were expected to indicate that each item provided unique information at varying levels of the trait. Good items should provide unique information and therefore discriminate well between various levels of the trait. Category response functions (CRF) were examined for each item to determine where the item functions in terms of the latent trait and how well the item discriminates across categories of responses. The associated boundary location parameters (*b-parameter*) for each item's categorical responses were examined to determine the point of overlap between each response category. The *b-parameter* indicates the trait level

required to score at or above a categorical threshold (Embretson & Reise, 2000). Good items were expected to provide unique information at each level of response and therefore have distinct boundaries. Although the *a-parameter* is not considered a discrimination parameter in the graded response model, it is an indication of "how quickly the expected item scores changes as a function of trait level" (Embretson & Reise, 2000, p. 103). Therefore, items with higher *a-parameters* should have CRF's that are more narrow and tall. Narrow and tall CRF's combined with the IIF is measure of discrimination.

It was expected that scoring higher on an item would indicate an increased presence of the latent trait being measured by the associated subscale. Additionally, items that discriminate well were expected to have tall and narrow CRF's and IIF's that indicate information is present at a range of the latent trait.

CHAPTER III

Results

For each subscale of the FFMQ, any item that was negatively scored was reverse coded prior to analysis. Internal consistency was assessed by way of coefficient alpha, item to subscale correlations, and alpha if the item were removed. The dimensionality of each facet was assessed using EFA with ML estimation. Finally, each subscale was examined using SGRM with EAP estimation. The TIF for each subscale can be seen in Figure A2. Graphic representation of strong and weak items under the SGRM can be seen in the example CRF and IIF in Figure A3 and Figure A4, respectively.

Depression Anxiety and Stress Scale

On average, participants were not experiencing symptom of depressions (M = 7.71, SD = 5.81), anxiety (M = 7.56, SD = 6.79), or stress (M = 12.94, SD = 8.74). Eighteen percent of participants reported moderate to severe levels of depression, 29% of participants reported moderate to severe levels of anxiety, and 25% reported moderate to severe levels of stress.

Five-Facet Mindfulness Scale

Reliability, validity, and EFA with ML estimation were used to assess the properties of the FFMQ, and the results were similar to that of prior research. Internal consistency of the FFMQ is within a satisfactory range for both clinical and research purposes, α = .86. Corrected item total correlations and alpha if the item were removed can be seen in Table A5. The average composite score was 125.42 (SD = 16.34), with a

reported range of scores of 74 to 179. Means and standard deviations for each item can also be seen in Table A5.

Incremental validity was assessed by examining the relationship between reported mindfulness, depression, anxiety, and age. As expected, mindfulness was negatively related to depression scores (r = -.41, p < .001), negatively related to anxiety (r = -.31, p < .001), and positively related to age (r = .17, p < .001).

An EFA with ML estimation was conducted to confirm the factor pattern that was expected based on prior research and theory. The KMO measure of sampling adequacy was .89, which is above the recommended value of .60. The Bartlett's test of sphericity was significant, χ^2 (741) = 9540.10, p < .001. The KMO and Bartlett's test results indicated that the assumptions of factor analysis were met. A clear five-facet pattern accounting for 51% of the variance in mindfulness emerged. The initial eigenvalue (6.83) accounted for 17% of the variance, the second eigenvalue (5.77) accounted for 15% percent of the variance, the third eigenvalue (2.92) accounted for 8% of the variance, the fourth eigenvalue (2.29) accounted for 6% of the variance, and finally the fifth eigenvalue (1.89) accounted for 5% of the variance. The pattern of loadings, communalities, and factor loadings can be seen in Table A5.

Finally, each of the five facets are considered to be distinct aspects of mindfulness and the weak to modest intercorrelations presented in Table A6 demonstrate the desired pattern of results, thus supporting the analysis of individuals facets (subscales) of the FFMQ.

Observing

The Observing scale consisted of 8 items with an average score of 26.54 (5.73). Internal consistency was within an acceptable range ($\alpha = .79$), with corrected item to subscale correlations ranging from $r = .42$ to $r = .61$. Corrected item to subscale correlations, expected change in coefficient alpha if the item was to be removed, and item level descriptive statistics can be seen in Table A7. When assessing the incremental validity of the Observing facet, a paradoxical pattern of correlations was revealed. Observing is positively related to depression ($r = .13$, $p = .001$), positively related to anxiety ($r = .26$, $p < .001$), and not significantly related to age ($r = .02$, $p = .586$).

An EFA using ML estimation revealed the KMO measure of sampling adequacy was .87 and the Bartlett's test of sphericity was significant, $\chi^2 (28) = 1066.03$, $p < .001$. When examining the factor structure of the Observing facet, a single factor with an eigenvalue of 3.28 that accounted for 41% of the variance was produced. Factor loadings for each item ranged from a low of .47 to a high of .71 and accounted for 22% to 50% of the variance in Observing. Communalities and factor loadings for each item can be seen in Table A7. Based on the totality of the EFA results, unidimensionality of the facet was assumed.

The IRT analysis of Observing using SGRM with EAP estimation converged in 27 loops. The TIF indicated that the scale produced the most information about the trait at a theta range of -2.5 to 2.0, with maximum information (4.17) provided at $\theta = -1.05$. Model fit for this scale was $\chi^2 / df = 1.25$, well below the cut-off of 3.0 indicating that the

data fit the model well. Additionally, based on item-to-model fit indices, all items fit the graded response model well.

Based CTT and IRT results, the top two performing items were 15 and 20, and the poorest were items 6 and 36. Item 15 accounted for 25% of the variance in Observing, had a corrected item-total correlation of r = .61, an *a-parameter* of 1.06 (0.07), distinct boundary locations, and an IIF that indicated that the item provided unique information at a theta range of -2.5 to 2.0. Item 20 performed modestly well, accounted for 15% of the variance, had a corrected item-total correlation of r = .53, an *a-parameter* of 0.88 (0.06), distinct boundary locations, and an IIF that indicated that the item provided unique information at a theta range of -3.0 to 2.0.

Item 6 accounted for 7% of the variance in Observing, had a corrected item-total correlation of r = .46, an *a-parameter* of 0.63 (0.04), poorly defined boundaries, and a relatively IIF that indicated little information was provided at any level of the latent trait. Likewise, item 36 accounted for 5% of the variance, had a corrected item-total correlation of r = .43, an *a-parameter* of 0.61 (0.04), poorly defined boundaries, and a flat IIF. Item parameters for the remaining items indicated that item 26 and 31 performed well with results similar to that of item 20, while items 1 and 36 performed poorly similar to item 6. Detailed item parameters for each item can be seen in Table A8.

Describing

The Describing scale also consisted of 8 items with an average score of 26.94 (5.91). Internal consistency was very good (α = .86), with corrected item to subscale

correlations ranging from $r = .46$ to $r = .69$. Corrected item to subscale correlations, expected change in coefficient alpha if the item was to be removed, and item level descriptive statistics can be seen in Table A7. When assessing the incremental validity of the Describing facet, the pattern of correlations was as expected. Describing was negatively related to depression ($r = -.19$, $p < .001$), negatively related to anxiety ($r = -.16$, $p < .001$), but not significantly related to age ($r = .08$, $p = .051$).

An EFA with ML estimation revealed that the KMO measure of sampling adequacy was .88 and the Bartlett's test of sphericity was significant, $\chi^2 (28) = 2022.97$, $p < .001$. When examining the factor structure of the Describing facet, a single factor with an eigenvalue of 4.08 that accounted for 51% of the variance was produced. Factor loadings for each item ranged from a low of .51 to a high of .75 and accounted for 26% to 56% of the variance in Describing. Communalities and factor loadings for each item can be seen in Table A7. Based on the totality of the EFA results, unidimensionality of the facet was assumed.

The IRT analysis of the Describing scale using SGRM with EAP estimation converged in 34 loops. The TIF indicated that the scale produced the most information about the latent trait at a theta range of -2.5 to 2.0, with maximum information (7.43) provided at $\theta = -1.10$. Model fit for this scale was $\chi^2/df = 1.31$ indicating that the data fit the model well. Additionally, based on item-to-model fit indices, all items fit the graded response model well.

Based on CTT results, items 2 and 37 were the top performing, while 22 and 32 were the weakest. Item 2 accounted for 29% of the variance in Describing and had a

corrected it-total correlation of r = .67, while item 37 accounted for 31% of the variance and had a corrected item-total correlation of r = .49. However, based on IRT results, the top two performing items were 2 and 16, and the poorest were items 22 and 32. Item 2 had an *a-parameter* of 1.248 (0.086), distinct boundary locations, and an IIF that indicated that the item provided unique information at a theta range of -3.5 to 2.5. Item 16 accounted for 23% of the variance, had a corrected item-total correlation of r = .65, an *a-parameter* of 1.148 (0.079), distinct boundary locations, and an IIF that indicated that the item provided unique information at a theta range of -3.5 to 2.5.

Item 22 accounted for 7% of the variance, had a corrected item-total correlation of r = .46, an *a-parameter* of 0.709 (0.048), poorly defined boundaries in response category 1 and 2, and a relatively flat IIF that indicated little information was provided at any level of the latent trait. Likewise, item 32 accounted for 9% of the variance, had a corrected item-total correlation of r = .49, an *a-parameter* of 0.710 (0.047), poorly defined boundaries in response category 1 and 2, and a flat IIF. Item parameters for the remaining items indicated that most items performed well with item parameters similar to that of item 2. Detailed item parameters for each item of the Describing facet can been seen in Table A8.

Acting with Awareness

The Acting with Awareness scale also consisted of 8 items with an average score of 25.94 (6.00). Internal consistency was very good (α = .87), with corrected item to subscale correlations ranging from r = .53 to r = .73. Corrected item to subscale correlations, expected change in coefficient alpha if the item was to be removed, and item

level descriptive statistics can be seen in Table A7. When assessing the incremental validity of the Acting with Awareness facet, the pattern of correlations was as expected. Acting with Awareness was negatively related to depression (r = -.37, p < .001), negatively related to anxiety (r = -.35, p < .001), and positively related to age (r = .14, p = .001).

An EFA using ML estimation revealed that the KMO measure of sampling adequacy was .87 and the Bartlett's test of sphericity was significant, χ^2 (28) = 2211.91, p < .001. When examining the factor structure of the Acting with Awareness facet, a single factor with an eigenvalue of 4.18 that accounted for 52% of the variance was produced. Factor loadings for each item ranged from a low of .52 to a high of .80 and accounted for 27% to 71% of the variance in Acting with Awareness. Communalities and factor loadings for each item can be seen in Table A7. Based on the totality of the EFA results, unidimensionality of the facet was assumed.

The IRT analysis of Acting with Awareness converged in 49 loops. The TIF indicated that the scale produced the most information about the latent trait at a theta range of -1.5 to 2.0, with maximum information (8.40) provided at θ = -0.35. Model fit for this facet was χ^2/df = 1.15 indicating that the data fit the model well. Additionally, based on item-to-model fit indices, all items fit the graded response model well.

Based on CTT results, items 5 and 13 were the best two performing items, while 23 and 34 were the weakest. Item 5 accounted for 48% of the variance in Acting with Awareness and had a corrected item-total correlation of r = .64, while item 13 accounted for 50% of the variance and had a corrected item-total correlation of r = .73. However,

based on IRT results, the top two performing items were 8 and 13, and the weakest were items 23 and 34. Item 8 accounted for 29% of the variance, had a corrected item-total correlation of r = .66, an *a-parameter* of 1.266 (0.087), distinct boundary locations, and an IIF that indicated that the item provided unique information at a theta range of -2.5 to 2.5. Item 13 had an *a-parameter* of 1.667 (0.119), distinct boundary locations, and an IIF that indicated that the item provided unique information at a theta range of -2.0 to 2.5.

Item 23 had an *a-parameter* of 0.740 (0.048), modestly defined boundaries, and a relatively flat IIF that indicated little information was provided at any level of the latent trait. Likewise, item 34 had an *a-parameter* of 0.735 (0.049), modestly defined boundaries, and a flat IIF. Item parameters for the remaining items indicated modest performance similar to that of item 23. Detailed item parameter for each item of the Acting with Awareness facet can been seen in Table A8.

Non-Judging of Inner Experiences

The Non-Judging of Inner Experiences scale also consisted of 8 items with an average score of 25.21 (6.91). Internal consistency was very good (α = .89), with corrected item to subscale correlations ranging from r = .60 to r = .74. Corrected item to subscale correlations, expected change in coefficient alpha if the item was to be removed, and item level descriptive statistics can be seen in Table A7. When assessing the incremental validity of the Non-Judging of Inner Experiences facet the pattern of correlations was as expected. Non-Judging of Inner Experiences was negatively related to depression (r = -.52, p < .001), negatively related to anxiety (r = -.44, p < .001), and positively related to age (r = .13, p = .002).

An EFA using ML estimation revealed that the KMO measure of sampling adequacy was .90 and the Bartlett's test of sphericity was significant, $\chi^2(28) = 2333.83$, $p < .001$. When examining the factor structure of the Non-Judging of Inner Experiences facet, a single factor with an eigenvalue of 4.48 that accounted for 56% of the variance was produced. Factor loadings for each item ranged from a low of .62 to a high of .83 and accounted for 38% to 69% of the variance in Non-Judging of Inner Experiences. Communalities and factor loadings for each item can be seen in Table A7. Based on the totality of the EFA results, unidimensionality of the facet was assumed.

The IRT analysis using SGRM with EAP estimation of Non-Judging of Inner Experiences converged in 40 loops. The TIF indicated that the scale produced the most information about the latent trait at a theta range of -2.0 to 1.5, with maximum information (9.42) provided at $\theta = -1.00$. Model fit for this scale was $\chi^2/df = 1.68$ indicating that the data fit the model well. Additionally, based on item-to-model fit indices, all items fit the graded response model well.

Based on CTT results, items 25 and 30 were the strongest performing, while items 3 and 39 were the weakest. Item 25 accounted for 41% of the variance in Non-Judging of Inner Experiences and had a corrected item-total correlation of $r = .74$, while item 30 accounted for 48% of the variance and had a corrected item-total correlation of $r = .76$. Item 3 accounted for 14% of the variance and had a corrected item-total correlation of $r = .60$, while item 39 accounted for 15% of the variance and had a corrected item-total correlation of $r = .60$.

However, based on IRT results, the top two performing items were 25 and 30, and the poorest were items 3 and 10. Item 25 had an *a-parameter* of 1.504 (0.106), distinct boundary locations, and an IIF that indicated that the item provided unique information at a theta range of -2.5 to 2.5. Item 30 had an *a-parameter* of 1.677 (0.120), distinct boundary locations, and an IIF that indicated that the item provided unique information at a theta range of -2.5 to 2.0.

Item 3 had an *a-parameter* of 0.878 (0.059), with modestly defined boundaries, and an IIF that indicates a small amount of information produced at a theta range of -2.5 to 3.0. Likewise, item 10 accounted for 17% of the variance, had a corrected item-total correlation of $r = .61$, an *a-parameter* of 0.930 (0.063), modestly defined boundaries, and produced a small amount of information at a theta range of -2.5 to 3.0. Item parameter for the remaining items indicate good performance with results similar to that of item 25. Detailed item parameters for each item of the Non-Judging of Inner Experiences facet can been seen in Table A8.

Non-Reactivity to Inner Experiences

The Non-Reactivity to Inner Experiences scale consisted of 7 items with an average score of 21.01 (4.31). Internal consistency was acceptable ($\alpha = .75$), with corrected item to subscale correlations ranging from $r = .38$ to $r = .56$. Corrected item to subscale correlations, expected change in coefficient alpha if the item was to be removed, and item level descriptive statistics can be seen in Table A7. When assessing the incremental validity of the Non-Reactivity to Inner Experiences subscale the pattern of correlations was as expected. Non-Reactivity to Inner Experiences was negatively

related to depression (r = -.12, p = .002), negatively related to anxiety (r = -.09, p = .017), and positively related to age (r = .11, p = .009).

An EFA with ML estimation revealed that the KMO measure of sampling adequacy was .82 and the Bartlett's test of sphericity was significant, χ^2 (21) = 798.63, p < .001. When examining the factor structure of the Non-Reactivity to Inner Experiences facet, a single factor with an eigenvalue of 2.83 that accounted for 14% of the variance was produced. Factor loadings for each item ranged from a low of .46 to a high of .67 and accounted for 21% to 45% of the variance in Non-Reactivity to Inner Experiences. Communalities and factor loadings for each item can be seen in Table A7. Based on the totality of the EFA results, unidimensionality of the facet was assumed.

The IRT analysis using SGRM with EAP estimation of Non-Reactivity to Inner Experiences converged in 28 loops. The TIF indicated that the scale produced the most information about the latent trait at a theta range of -2.5 to 3.0, with maximum information (3.19) provided at θ = -0.90. Model fit for this scale was χ^2/df = 1.92 indicating that the data fit the model well. Additionally, based on item-to-model fit indices, all items fit the graded response model well.

Based on CTT results, the strongest items were 24 and 33, while the weakest were 21 and 29. Item 24 accounted for 9% of the variance in Non-Reactivity to Inner Experiences and had a corrected item-total correlation of r = .35, while item 33 accounted for 12% of the variance and had a corrected item-total correlation of r = .51. Item 21 accounted for 5% of the variance and had a corrected item-total correlation of r = .45,

while item 29 accounted for 4% of the variance and had a corrected item-total correlation of r = .56.

Based on IRT results, the top two performing items were 29 and 33, and the poorest were items 4 and 24. Item 29 had an *a-parameter* of 1.008 (0.067), relatively distinct boundary locations, and an IIF that indicated that the item provided a small amount of information at a theta range of -3.0 to 3.0. Item 33 accounted for 12% of the variance, had a corrected item-total correlation of r = .51, an *a-parameter* of 0.873 (0.058), modestly defined boundary locations, and an IIF that indicated that the item provided a small amount of information at a theta range of -2.0 to 3.0.

Item 4 accounted for 6% of the variance, had a corrected item-total correlation of r = .44, an *a-parameter* of 0.585 (0.038), poorly defined boundaries in all response categories, and a relatively flat IIF that indicated little information was provided at any level of the trait. Likewise, item 24 had an *a-parameter* of 0.577 (0.038), poorly defined boundaries in response categories 2 through 5, and a flat IIF. Item parameters for the remaining items indicated poor performance similar to that of item 4. Detailed item parameters for each item of the Non-Reactivity to Inner Experiences can been seen in Table A8.

CHAPTER IV

Discussion

The purpose of the current study was to conduct an item level analysis of the Five-Facet Mindfulness Questionnaire, a measure commonly used in mindfulness-based clinical practice and research. Results based on CTT properties were similar to those found in prior studies. Describing, Acting with Awareness, and Non-Judging of Inner Experiences facets produced the most reliable results as indicated by coefficient alpha and communalities. For each of these three facets, the majority of items accounted for a range of 20% to 50% of the variance in the scale scores. Based on low communalities and little to no change expected in reliability upon removal, item 22 and 32 of the Describing facets are good candidates for removal or revision. Similarly, items 23, 28, and 34 of the Acting with Awareness facet are good candidates for removal or revision. Based on CTT results, all items of the Non-Judging of Inner Experiences performed well.

The Observing and Non-Reactivity to Inner Experiences facets provided much weaker results. Both scales produced a coefficient alpha slightly above the .70 cut-off. However, with the exception of item 15 on the observing scale, none of the items accounted for anymore than 15% of the variance in scale scores. Based on CTT results, the Observing facet and Non-Reactivity to Inner Experiences facet need major revisions.

An assessment of the correlation between each facet, age, and measures of depression or anxiety revealed the expected relationship in all facets except Observing. Although this is a paradoxical finding, it was not entirely unexpected. Similar results

have been found in non-meditating samples (Baer et al., 2006). One possible explanation for this finding is that focusing attention on the self can lead to maladaptive rumination and has been shown to be related to negative emotions such as those seen in depression and anxiety (Mor & Winquist, 2002). Mindfulness training involves learning to observe inner experiences in a non-reactive and non-judgmental fashion. Thus, the practice of mindfulness is expected to lead to a reduction in maladaptive rumination and the ability to observe negative thoughts and feelings in a way that is nonjudgmental and produces acceptance (Segal, Williams, & Teasdale, 2002; Watkins & Teasdale, 2004). Although, this is a plausible explanation, the relationship reversal between Observing and maladaptive constructs is also an indication of problems within the Observing scale. If the FFMQ is to be used as a research tool measuring change in mindfulness before and after mindfulness training, it is necessary for the scale to be sample invariant, thus producing further evidence for major revision of the Observing scale.

Upon confirming that each facet was comprised of a single dominant factor, the TIF for each scale was inspected. It was hypothesized that the each scale should provide information across a range of mindfulness. No facet produced good information in the far upper ($\theta > 2.0$) or lower ($\theta < -2.5$) range of the latent trait, perhaps indicating the need for a more sensitive measure. However, Describing, Acting with Awareness, and Non-Judging of Inner Experiences performed quite well at a theta range of -2.5 to 2.0. It should be noted that although item 2 on the Describing scale appeared to produce good information and discrimination, its fit to the graded response model was questionable ($\chi^2/df = 2.45$). Although the item to model fit for item 2 was below the cut-off

recommended by Chernyshenko, Stark, Chan, Dragsow, and Williams (2001), it was much worse than that of the other items in the Describing scale. A similar finding was apparent for item 25 ($\chi^2/df = 2.60$) on the Non-Judging of Inner Experiences scale.

Observing and Non-Reactivity to Inner Experiences also performed at a theta range of -2.5 to 2.0, but produced little information regardless of the level of the latent trait. Three items on the Non-Reactivity to Inner Experiences (19, 29, and 33) produced questionable item to model fit indices. This pattern of results suggests that the Observing and Non-Reactivity to Inner Experiences scales are not reliable measures, and none of the scales are reliable in either the highest or lowest ranges of latent trait. Although not ideal, the range of information provided is potentially sufficient for measuring mindfulness in a therapeutic or research setting where scores at the extreme ends of the scale would not be expected.

An examination of the performance of each item in the context of the *a-parameter*, CRF and IIF revealed a pattern of results similar to that of the CTT analyses, with a great deal of instability in Non-Reactivity to Inner Experiences. For the Observing and Non-Reactivity to Inner Experience scales, most items performed poorly as indicated by low *a-parameters*, short and wide CRF's, and low flat IIF's. Items on the Observing scale measured one's tendency to observe thoughts, emotion, and sensations, and only one item (15) performed within an acceptable range of measurement. The Non-Reactivity to Inner Experiences scale measured one's tendency to react to negative thoughts, images, and emotions. Items 29 and 33 on this scale performed slightly better than other non-reactivity items under IRT, but under CTT item 33 was the strongest while item 33 was

the weakest, indicating instability in the scale. No item on the Non-Reactivity to Inner Experiences scale produced an adequate amount of unique information under IRT and performed equally poor under CTT. Additionally, items on the Observing and Non-Reactivity to Inner Experiences scales produced CRF's that were relatively flat and overlapping, little to no information at any level of the latent trait, and low communalities. This pattern of results provides further evidence for major revisions in both measures.

The remaining scales performed with much greater precision. Items on the Describing scale measured one's ability or tendency to express thought, feelings, sensations, and experiences in words. With the exception of items 22 and 32, the items on the Describing scale produced CRF's with tall and narrow peaks, IIF's indicating information was provided at wide range of the latent trait, and commonalties that indicated each item accounted for a minimum of 20% of the scale score. Item 22 and 32 provided little information and appeared to perform best at lower levels of the latent trait, indicating the need for revision.

The items on the Acting with Awareness scale were a measure of attentional regulation. All items performed well with the weakest items being 18, 23, 28, and 34. These four items had reduced discrimination as indicated by IIF (low and flat). The CRF for these items indicated less distinct boundaries, but the CRF and IIF combined indicated that like items 22 and 32, these items performed better at lower levels of the latent trait. The Non-Judging of Inner Experiences scale measured one's tendency to be self-critical. With the exception of item 3, each item performed moderately well. Each

item produced information at a range of the trait, had CRF's with distinct categorical boundaries, and had acceptable communalities.

The IRT analyses of Describing, Acting with Awareness, and Non-Judging of Inner Experiences indicate that the items of these scales generate moderate to good discrimination and adequate information about their underlying trait. Likewise, under CTT these scales demonstrate strong reliability and perform well as a whole. The CTT and IRT results produced similar findings and indicated that a handful of items on these three scales need revision in order to create a more sensitive measure of mindfulness.

However, the same cannot be said for Observing and Non-Reactivity to Inner Experiences. The totality of CTT analyses and IRT analyses suggests that the Observing and Non-Reactivity to Inner Experiences scales need major revision. Both scales produced low levels of reliability and have items that provide little information and discriminate poorly across levels of the latent trait. Observation and non-reactivity are both central tenets of mindfulness, but have not been optimally operationalized in the FFMQ.

The FFMQ as a whole performs well in terms of coefficient alpha, but varies from sample to sample in terms of data to model fit in factor analysis (Fernandex, Wood, Stein & Rossi, 2010; Bohlmeijer, ten Klooster, Fledderus, Veehof, & Baer, 2011; Veehof, ten Klooster, Taal, Westerhof, & Bohlmeijer, 2011; Hou, Wong, Lo, Mak, & Ma 2013). The current study indicates that three of the five facets of mindfulness are well represented with strong items. However, the weak performance of items from the Observing facet and the Non-Reactivity to Inner Experiences facet shed light on the

erratic results when attempting to factor analyze the construct of mindfulness without using item parceling. Although the FFMQ as a whole appears to be sufficient for mindfulness-based clinical purposes, more work is needed. Much has been learned about the physiological and psychological benefits of mindfulness training, but little is known about the mechanism of action behind that change. In order to elucidate potential mechanisms of action behind the changes brought about by mindfulness meditation, the theory needs to be optimally operationalized and better measures need to be created. Future research should investigate changes to the Observing and Non-Reactivity to Inner Experiences scales.

A major limitation of the current study was that a relatively small sample of participants with meditation experience was analyzed. An additional weakness was the lack of measurement for length of experience with meditation and type of meditation practiced. Therefore, future research should also focus on incorporating a larger sample of meditating individuals who have been trained in the practice of mindfulness. Additionally, given the pattern of correlations seen in the Observing facet, a combined sample incorporating larger number of meditating participants would allow for measurement of differential item or test functioning between meditating and non-meditating participants on all facets of mindfulness.

REFERENCES

Assessment Systems Corporation. (1996). *User's manual for the XCALIBRE marginal maximum-likelihood estimation program*. St. Paul, MN: Assessment Systems Corp.

Baer, R. (2011). Measuring Mindfulness. *Contemporary Buddhism: An Interdisciplinary Journal, 12*:1, 241-261. doi**:**10.1080/14639947.2011.564842

Baer, R., Smith, G., & Allen, K. (2004). Assessment of mindfulness by self-report: The Kentucky Inventory of Mindfulness Skills. *Assessment, 11*(3), 191-206. doi: 10.1177/1073191104268029

Baer, R., Smith, G., Hopkins, J., Krietemeyer, J., & Toney, L. (2006). Using self-report assessment methods to explore facets of mindfulness. *Assessment, 13*(1), 27-45. doi: 10.1177/1073191105283504

Baer, R., Smith, G., Lykins, E., Button, D., Krietemeyer, J., Sauer, S., Walsh, E., Duggan, D., & Williams, J. (2008). Construct validity of the five facet mindfulness questionnaire in meditating and nonmeditating samples. *Assessment, 15*(3), 329-342. doi: 10.1177/1073191107313003

Baker, A. (2001). *The basics of item response theory. (2nd ed.)*. ERIC Clearinghouse on Assessment and Evaluation.

Bandalos, D. L. (2002). The effects of item parceling on goodness-of-fit and parameter estimate bias in structural equation modeling. *Structural Equation Modeling, 9*(1), 78-102. doi: 10.1207/S15328007SEM0901_5

Bartlett, M. S. (1954). A note on multiplying factors for various chi-squared approximations. *Journal of the Royal Statistical Society, Series B*(16), 296–298.

Bergomi, C., Tschacher, W., & Kupper, Z. (2013). The assessment of mindfulness with self-report measures: Existing scales and open issues. *Mindfulness, 4*(3), 191-202. doi: 10.1007/s12671-012-0110-9

Birnbaum, A. (1968). Some latent trait models and their use in inferring an examinee's ability. In Lord, F.M. & Novick, M.R. (Eds.), *Statistical theories of mental test scores*. Reading, MA: Addison–Wesley.

Bohlmeijer, E. T., ten Klooster, P. M., Fledderus, M., Veehof, M. M., & Baer, R. A. (2011). Psychometric properties of the Five Facet Mindfulness Questionnaire in depressed adults and development of a short-form. *Assessment, 18*(3), 308-320. doi: 10.1007/s10067-011-1690-9

Brown, K., & Ryan, R. (2003). The benefit of being present: Mindfulness and its role in psychological well-being. *Journal of Personality and Social Psychology, 84,* 822-848. doi: 10.1037/0022-3514.84.4.822

Brown, K., Ryan, R., & Creswell, J. W. (2007). Mindfulness: Theoretical foundations and evidence for its salutary effects. *Psychological Inquiry, 18,* 211-237. doi:10.1080/10478400701598298

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 of Meditation and Meditation Research*, 1, 11-34.

Cattell, R. B. (1966). The scree test for the number of factors. *Multivariate Behavioral Research, 1,* 629-637. doi: 10.1207/s15327906mbr0102_10

Carmody, J., & Baer, R.A. (2008). Relationships between mindfulness practice and levels of mindfulness, medical and psychological symptoms and well-being in a mindfulness-based stress reduction program. *Journal of Behavioral Medicine, 31*, 23-33. doi: 10.1007/s10865-007-9130-7

Chadwich, P., Hember, M., Smyes, J., Peters, E., Kuipers, E., & Dagnan, D. (2008). Responding mindfully to unpleasant thoughts and images: Reliability and validity of the Southampton mindfulness questionnaire. *British Journal of Clinical Psychology, 47*(4), 451-455. doi:10.1348/014466508X314891

Christopher, M. S., Neuser, N. J., Michael, P. G., & Baitmangalkar, A. (2012). Exploring the psychometric qualities of the Five Facet Mindfulness Questionnaire. *Mindfulness, 3*(2), 124-131. doi: 10.1007/s12671-011-0086-x

Comrey, A. L., & Lee, H. B. (1992). *A first course in factor analysis* (2nd ed.). Hillsdale, NJ: Lawrence Erlbaum Associates.

Cronbach, L. J. (1951). Coefficient alpha and the internal structure of tests. Psychometrika. 16, 297-334. doi: 10.1007/BF02310555

Dragsow, F., & Parsons, C. K. (1983). Application of unidimensional item response theory models to multidimensional data. *Applied Psychological Measurement, 7*(2), 189-199. doi: 10.1177/014662168300700207

Embretson, S. E., & Reise, S. (2000). *Item response theory for psychologists*. Mahwah, NJ: Erlbaum Publishers.

Fernandez, A. C., Wood, M. D., Stein, L. A. R., & Rossi, J. S. (2010). Measuring mindfulness and examining its relationship with alcohol and negative consequences. *Psychology of Addiction, 24*(4), 608-616. doi:10.1037/a0021742

Floyd, F. J., & Widamon, K. F. (1995). Factor analysis in the development and refinement of clinical assessment instruments. *Psychological Assessment, 3*, 286-299. doi: 10.1037//1040-3590.7.3.286

Hayes, A. M., & Feldman, G. C. (2004). Clarifying the construct of mindfulness in the context of emotion regulation and the process of change in therapy. *Clinical Psychology: Science and Practice, 11,* 225-262. doi:10.1093/clipsy.bph080

Hou, J., Wong, S. Y., Lo, H. H., Mak, W. W., & Ma, H. S. (2013). Validation of a Chinese version of the five facet mindfulness questionnaire in Hong Kong and development of a short-form. *Assessment.* Retrieved from http://asm.sagepub.com/content/early/2013/04/16/1073191113485121

IBM Corp. (2011). IBM SPSS Statistics for Windows, Version 20.0. Armonk, NY: IBM Corp.

Joreskog, K. G. (1969). A general approach to confirmatory maximum-likelihood factor analysis. *Psychometrika, 34,* 183-202. doi: 10.1007/BF02289343

Kabat-Zinn, J. (1990). *Full catastrophe living: The program of the Stress Reduction Clinic at the University of Massachusetts Medical Center.* New York: Delta.

Kabat-Zinn, J. (1994). *Wherever you go, there you are: Mindfulness meditation in everyday life.* New York: Hyperion Books.

Kabat-Zinn, J. (2000). Indra's net at work: The mainstreaming of Dharma practice in society. In G. Watson & S. Batchelor (Eds.), *The psychology of awakening: Buddhism, science, and our day-to-day lives* (pp. 225-249). North Beach, ME: Weiser.

Kaiser, H. F. (1960). The application of electronic computers to factor analysis. *Educational and Psychological Measurement, 20,* 141-151. doi: 10.1177/001316446002000116

Kaiser, H.F. (1974). An index of factorial simplicity. *Psychometrika*, 39, 31-36. doi: 10.1007/BF02291575

Leary, M. (2004). *The curse of the self: Self-awareness, egotism and quality of human life.* New York, NY: Oxford University Press.

Linehan, M. M. (1993). *Cognitive behavioral therapy of borderline personality disorder*, Guilford Press, New York.

Little, R. J. A. (1998) A test of missing completely at random for multivariate data with missing values. *Journal of the American Statistical Association,83*, 1198 – 1202. doi: 10.2307/2290157

Lord, F. M., & Novick, M. R. (1968). *Statistical theories of mental test scores*. Reading, Mass: Addison-Wesley.

Marsh, H.W., Ellis, L., Parada, L., Richards, G., & Heubeck, B.G. (2005). A short version of the Self Description Questionnaire II: Operationalizing criteria for short-form evaluation with new applications of confirmatory factor analyses. *Psychological Assessment, 17,* 81–102. doi: 10.1037/1040-3590.17.1.81

Mor, N., & Winquist, J. (2002). Self-focused attention and negative affect: A meta-analysis. *Psychological Bulletin, 128,* 638-662. doi: 10.1037//0033-2909.128.4.638

Nunnally, J. C. (1978). *Psychometric theory* (2nd ed.). New York: McGraw-Hill.

Reckase, M. D. (1979). Unifactor latent trait models applied to multifactor tests: Results and implications. *Journal of Educational and Behavioral Statistics, 4*(3), 207-230. doi:10.3102/10769986004003207

Reckase, M. D., & McKinely, R. L. (1991). The discriminating power of items that measure more than one dimension. *Applied Psychological Measurement, 15*(4), 361-373. doi:10.1177/014662169101500407

Rubin, D.B. (1987). *Multiple imputation for nonresponse in surveys*. John Wiley & Sons, New York.

Samejima, F. (1969). Estimation of latent ability using a response pattern of graded scores. *Psychometrika Monograph Supplement, 34*(4, Pt. 2). doi: 10.1007/BF02290599

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. doi: 10.1037/0022-006X.76.3.524

Shapiro, S., Oman, D., Thoresen, C., Plante, T., & Flinders, T. (2008). Cultivating mindfulness: Effects on well-being, *Journal of Clinical Psychology, 64*(7), 840-862. doi:10.1002/jclp.20491

Smith, R. M., & Miao, C. Y. (1994). Assessing unidimensionality for Rasch measurement. In M. Wilson (Ed.): *Objective Measurement: Theory into Practice. Volume 2.* Greenwich: Ablex.

Tabachnick, B. G., & Fidell, L. S. (2013). *Using Multivariate Statistics , 6th ed.* Boston: Allyn and Bacon.

Teasdale, J. D., Moore, R. G., Hayhurst, H., Pope, M., Williams, S., & Segal, Z. V. (2002). Metacognitive awareness and prevention of relapse in depression: Empirical evidence. *Journal of Consulting and Clinical Psychology, 70,* 278-287. doi: 10.1037//0022-006X.70.2.275

Teasdale, J., Segal, Z., Williams, J., Ridgeway, V., Soulsby, J., & Lau, M. (2000). Prevention of relapse/recurrence in major depression by mindfulness-based cognitive therapy. *Journal of Counseling and Clinical Psychology, 68*, 615-625. doi:10.1037//0022-006X.68.4.615

Turner, R. (2000). Naturalistic evaluation of dialectical behavior therapy-oriented treatment for borderline personality disorder. *Cognitive and Behavior Practice, 7*, 413-419. doi:10.1016/S1077-7229(00)80052-8

Uebersax, J. S. (1993). Statistical modeling of expert ratings on medical treatment appropriateness. *Journal of the American Statistical Association, 88,* 421-427. doi: 10.2307/2290320

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

Walach, H., Buchheld, N., Buttenmüller, V., Kleinknecht, N., & Schmidt, S. (2006). Measuring mindfulness: The Freiburg Mindfulness Inventory (FMI). *Personality and Individual Differences, 40*, 1543-1555.

Watkins, E., & Teasdale, J. D. (2004). Adaptive and maladaptive self-focus in depression. *Journal of Affective Disorders, 82,* 1-8. doi:10.1016/j.jad.2003.10.006

APPENDICES

APPENDIX A: TABLES AND FIGURES

Table A1

Published Mindfulness Questionnaires and Example Items

Freiburg Mindfulness Inventory
I am open to the experience of the present moment.
I sense my body, whether eating, cooking, cleaning, or talking.
When I notice an absence of mind I gently return to the experience of the here and now.
Mindful Attention Awareness Scale
I find myself doing things without paying attention. (R)
I break or spill things because of carelessness, not paying attention, or thinking of soothing else. (R)
It seems I am "running on automatic" without much awareness of what I'm doing. (R)
Kentucky Inventory of Mindfulness Skills
When I'm walking, I deliberately notice the sensations of my body moving.
I'm good at finding the words to describe my feelings.
When I do things, my mind wanders off and I'm easily distracted. (R)
I tell myself that I shouldn't be feeling the way I'm feeling. (R)
Cognitive and Affective Mindfulness Scale - Revised
I am able to focus on the present moment.
I am preoccupied by the past. (R)
I am able to accept the thoughts and feelings I have.
Southampton Mindfulness Questionnaire
When I have distressing thoughts of images, I am able to just notice them without reacting.
When I have distressing thoughts or images, I judge the thought or image as good or bad. (R)
When I have distressing thoughts of images in my mind I try and push them away. (R)
Five Facet Mindfulness Questionnaire
(This is a composite of the preceding five questionnaires and includes items from each.)

Note: R = reverse-scored items. Adapted from "Measuring Mindfulness", by R. A. Baer, (2011), *Contemporary Buddhism: An Interdisciplinary Journal, 12*:1, 241-261.

Table A2

Reliability, Intercorrelation, and Model Fit Indices of the FFMQ

	α	Observe	Describe	ActAware	NonJudge	χ^2	CFI	RMSEA
Baer et al. (2006) - Non-Meditating Sample								
Model Fit -Hierarchical						207.75**	.93[1]	.06[1]
Model Fit - Five-Factor Correlated						146.68**	.96[1]	.06[1]
Model Fit - One Factor						1113.78**	.43[1]	.21[1]
Observe	.83							
Describe	.91	.26**						
ActAware	.87	.15**	.30**					
NonJudge	.87	-.07**	.21**	.34**				
NonReact	.75	.16**	.22**	.33**	.34**			
Baer et al. (2008) - Meditating Sample								
Model Fit - Hierarchical						-	.97[1]	.06[1]
Model Fit - Five-Factor Correlated						-	.97[1]	.07[1]
Observe	-							
Describe	-	.40**						
ActAware	-	.43**	.32**					
NonJudge	-	.49**	.38**	.39**				
NonReact	-	.56**	.39**	.49**	.52**			
Christopher et al. (2012) - Mixed Sample								
Model Fit - Hierarchical						1672.45**	.97	.06
Observe	.84							
Describe	.91	.31**						
ActAware	.90	.26**	.43**					
NonJudge	.93	.13**	.40**	.54**				
NonReact	.86	.40**	.41**	.38**	.44**			
Bohlmeijer et al. (2011) - Non-Meditating Clinical Sample								
Model Fit - Hierarchical						2186.91**	.91	.08
Model Fit - Five-Factor Correlated						2126.90**	.91	.07
Observe	.78							
Describe	.91	.41**						
ActAware	.86	.29**	.30**					
NonJudge	.86	.02	.07	.23**				
NonReact	.73	.15**	.15**	.20**	.35**			
Fernandez et al. (2010) - Non-Meditating Clinical Sample								
Model Fit - Five-Factor Correlated						335.84**	.90[1]	.10[1]
Observe	-							
Describe	-	.38**						
ActAware	-	-.19**	.11					
NonJudge	-	-.35**	-.09	.42**				
NonReact	-	.37**	.35**	-.07	-.07			

Note: FFMQ = Five Facet Mindfulness Qestionnaire, FFMQ-SF = Five Facet Mindfulness Questionnaire Short-From, ActAware = acting with awareness, NonJudge = nonjudging of inner experiences, NonReact = nonreactivity to inner experiences.

1= parceled model.

***p < .01.*

Table A3

Items Retained in the Final Version of the Five-Facet Mindfulness Questionnaire

Item #	Origin	Item
1	KIMS	When I'm walking, I deliberately notice the sensations of my body moving.
2	KIMS	I'm good at finding words to describe my feelings.
3	KIMS	I criticize myself for having irrational or inappropriate emotions. (R)
4	FMI	I perceive my feelings and emotions without having to react to them.
5	KIMS	When I do things, my mind wanders off and I'm easily distracted. (R)
6	KIMS	When I take a shower or bath, I stay alert to the sensations of water on my body.
7	KIMS	I can easily put my beliefs, opinions, and expectations into words.
8	KIMS	I don't pay attention to what I'm doing because I'm daydreaming, worrying, or otherwise distracted. (R)
9	FMI	I watch my feelings without getting lost in them.
10		I tell myself I shouldn't be feeling the way I'm feeling. (R)
11	KIMS	I notice how foods and drinks affect my thoughts, bodily sensations, and emotions.
12	KIMS	It's hard for me to find the words to describe what I'm thinking. (R)
13	CAMS	I am easily distracted. (R)
14	KIMS	I believe some of my thoughts are abnormal or bad and I shouldn't think that way. (R)
15	KIMS	I pay attention to sensations, such as the wind in my hair or sun on my face.
16	KIMS	I have trouble thinking of the right words to express how I feel about things. (R)
17	KIMS	I make judgments about whether my thoughts are good or bad. (R)
18	MAAS	I find it difficult to stay focused on what's happening in the present. (R)
19	SMQ	When I have distressing thoughts or images, I "step back" and am aware of the thought or image without getting taken over by it.
20	KIMS	I pay attention to sounds, such as clocks ticking, birds chirping, or cars passing.
21	FMI	In difficult situations, I can pause without immediately reacting.
22	KIMS	When I have a sensation in my body, it's difficult for me to describe it because I can't find the right words. (R)
23	MAAS	It seems I am "running on automatic" without much awareness of what I'm doing. (R)
24	SMQ	When I have distressing thoughts or images, I feel calm soon after.
25	KIMS	I tell myself that I shouldn't be thinking the way I'm thinking. (R)
26	KIMS	I notice the smells and aromas of things.
27	KIMS	Even when I'm feeling terribly upset, I can find a way to put it into words.
28	MAAS	I rush through activities without being really attentive to them. (R)
29	SMQ	When I have distressing thoughts or images I am able just to notice them without reacting.
30	KIMS	I think some of my emotions are bad or inappropriate and I shouldn't feel them. (R)
31	KIMS	I notice visual elements in art or nature, such as colors, shapes, textures, or patterns of light and shadow.
32	KIMS	My natural tendency is to put my experiences into words.
33	SMQ	When I have distressing thoughts or images, I just notice them and let them go.
34	MAAS	I do jobs or tasks automatically without being aware of what I'm doing. (R)
35	SMQ	When I have distressing thoughts or images, I judge myself as good or bad, depending what the thought/image is about. (R)
36	KIMS	I pay attention to how my emotions affect my thoughts and behavior.
37	CAMS	I can usually describe how I feel at the moment in considerable detail.
38	MAAS	I find myself doing things without paying attention. (R)
39	KIMS	I disapprove of myself when I have irrational ideas. (R)

Note: R = reverse-scored item; Origin = scale from with the item was originally derived; CAMS = Cognitive and Affective Mindfulness Scale; FMI = Freiburg Mindfulness Inventory; KIMS = Kentucky Inventory of Mindfulness Skills; MAAS = Mindful Attention and Awareness Scale; SMQ = Southampton Mindfulness Questionnaire.

Table A4
Interpretation of the Discrimination Parameter in the Two and Three-Parameter Logistic Model

Interpretation	*a-parameter range*
No Discrimination	.01 -.34
Low	.35 - .64
Moderate	.65 - 1.34
High	1.35 - 1.69
Very High	> 1.70
Perfect	+ infinity

Table A5

Item Level and Test Level Statistics:, Descriptive Statistics, Item-Total Correlations, and Reliability for the FFMQ

Item	Facet	*M*	*SD*	*Corrected Item-Total Corr.*	*α If Item Removed*	*SMC / Communalities*	*Factor Loading*
1	OB	2.62	1.13	.17	.86	.29	.56
2	DS	3.35	1.03	.45	.85	.52	.68
3	NJ	2.92	1.20	.27	.85	.42	.56
4	NR	3.10	0.98	.16	.86	.27	.45
5	AA	2.71	1.17	.38	.85	.65	.80
6	OB	3.10	1.21	.23	.85	.25	.48
7	DS	3.52	1.07	.46	.86	.47	.62
8	AA	3.30	1.11	.47	.85	.55	.66
9	NR	3.11	0.95	.33	.85	.23	.45
10	NJ	2.92	1.19	.33	.85	.42	.61
11	OB	2.89	1.27	.21	.86	.24	.46
12	DS	3.40	1.05	.47	.85	.64	.83
13	AA	2.84	1.14	.45	.85	.73	.83
14	NJ	3.35	1.13	.46	.85	.59	.74
15	OB	3.35	1.15	.19	.86	.44	.76
16	DS	3.42	1.06	.47	.85	.64	.85
17	NJ	2.84	1.12	.18	.86	.47	.66
18	AA	3.55	1.02	.47	.85	.50	.56
19	NR	3.12	1.07	.35	.85	.33	.46
20	OB	3.61	1.08	.17	.86	.36	.61
21	NR	3.24	0.92	.28	.85	.26	.44
22	DS	3.56	0.89	.42	.85	.37	.54
23	AA	3.47	1.03	.38	.85	.34	.53
24	NR	2.59	1.01	.27	.85	.23	.46
25	NJ	3.23	1.08	.40	.85	.64	.83
26	OB	3.83	0.97	.32	.85	.38	.60
27	DS	3.29	1.07	.54	.85	.48	.55
28	AA	3.52	0.86	.35	.85	.35	.57
29	NR	2.97	0.97	.28	.85	.43	.67
30	NJ	3.40	1.17	.42	.85	.71	.89
31	OB	3.50	1.16	.26	.85	.36	.60
32	DS	3.16	1.08	.38	.85	.39	.44
33	NR	2.88	0.91	.30	.85	.47	.75
34	AA	3.37	0.96	.23	.86	.37	.66
35	NJ	3.29	1.18	.35	.85	.48	.68
36	OB	3.62	0.98	.25	.85	.33	.38
37	DS	3.26	1.04	.52	.85	.60	.69
38	AA	3.19	0.99	.33	.85	.46	.72
39	NJ	3.26	1.21	.28	.85	.40	.61

Note: FFMQ = Five Facet Mindfulness Questionnaire, OB = Observing, DS = Describing, AA = Acting with Awareness, NJ = Non-Judging on Inner Experiences, NR = Non-Reactivity to Inner Experiences.

Table A6

Intercorrelation of the FFMQ Facets

Facet	OB	DS	AA	NJ
DS	.31**			
AA	-.06	.30**		
NJ	-.20**	.15**	.34**	
NR	.37**	.29**	.05	-.01

Note: FFMQ = Five Facet Mindfulness Questionnaire, FFMQ-SF = Five Facet Mindfulness Questionnaire Short-From, OB = observing, DS = describing, AA = acting with awareness, NJ = nonjudging of inner experiences, NR = nonreactivity to inner experiences.

*$^{**}p < .001$.

Table A7

Item Level and Test Level Statistics: Descriptive Statistics, Item-Total Correlations, and Reliability for Each Subscale.

Facet / Item	α	*M*	*SD*	*Corrected Item-Subscale Corr.*	*α If Item Removed*	*SMC / Communalities*	*Factor Loading*
OB	.79						
1		2.62	1.13	.48	.77	.29	.54
6		3.10	1.21	.46	.77	.27	.52
11		2.89	1.27	.46	.76	.26	.51
15		3.35	1.15	.61	.75	.50	.71
20		3.61	1.08	.53	.76	.39	.62
26		3.83	0.97	.52	.77	.35	.59
31		3.50	1.16	.50	.77	.34	.59
36		3.62	0.98	.43	.78	.22	.47
DS	.86						
2		3.35	1.03	.67	.84	.54	.73
7		3.52	1.07	.63	.84	.47	.68
12		3.40	1.05	.64	.84	.47	.69
16		3.42	1.06	.65	.84	.48	.70
22		3.56	0.89	.46	.86	.26	.51
27		3.29	1.07	.63	.84	.47	.69
32		3.16	1.08	.49	.86	.30	.55
37		3.26	1.04	.49	.84	.56	.75
AA	.87						
5		2.71	1.17	.69	.85	.64	.80
8		3.30	1.11	.66	.85	.54	.74
13		2.84	1.14	.73	.84	.71	.84
18		3.55	1.02	.62	.85	.46	.67
23		3.47	1.03	.55	.86	.29	.54
28		3.52	0.86	.56	.86	.31	.56
34		3.37	0.96	.54	.86	.27	.52
38		3.19	0.99	.63	.85	.39	.63
NJ	.89						
3		2.92	1.20	.60	.88	.38	.62
10		2.92	1.19	.61	.88	.41	.64
14		3.35	1.13	.68	.87	.44	.66
17		2.84	1.12	.62	.88	.58	.76
25		3.23	1.08	.74	.86	.64	.80
30		3.40	1.17	.76	.86	.69	.83
35		3.29	1.18	.65	.87	.47	.69
39		3.26	1.21	.60	.88	.39	.62
NR	.75						
4		3.10	0.98	.44	.73	.25	.50
9		3.11	0.95	.45	.72	.26	.51
19		3.12	1.07	.49	.72	.25	.58
21		3.24	0.92	.45	.72	.22	.52
24		2.59	1.01	.38	.74	.30	.46
29		2.97	0.97	.56	.70	.17	.67
33		2.88	0.91	.51	.70	.34	.61

Note: FFMQ = Five Facet Mindfulness Questionnaire, OB = observing, DS = describing, AA = acting with awareness, NJ = nonjudging of inner experiences, NR = nonreactivity to inner experiences.

Table A8

Item Fit Indices, a-Parameter, and Boundary Location Parameters for the best two items of each facet of the FFMQ.

Item Facet	X^2/df	*a-parameter*	*b1*	*b2*	*b3*	*b1*
OB						
1	1.22	0.667 (0.046)	-1.465 (0.079)	-0.368 (0.062)	1.439 (0.081)	3.090 (0.156)
6	0.91	0.631 (0.042)	-2.529 (0.125)	-0.753 (0.068)	0.463 (0.065)	1.971 (0.101)
11	1.15	0.622 (0.042)	-1.759 (0.089)	-0.576 (0.064)	0.773 (0.069)	2.274 (0.111)
15	1.53	1.061 (0.072)	-2.066 (0.091)	-1.088 (0.054)	0.023 (0.047)	1.386 (0.070)
20	1.29	0.882 (0.061)	-2.611 (0.120)	-1.663 (0.072)	-0.312 (0.055)	1.159 (0.070)
26	1.69	0.834 (0.059)	-3.326 (0.184)	-2.189 (0.094)	-0.777 (0.061)	0.986 (0.070)
31	1.13	0.787 (0.054)	-2.545 (0.117)	-1.408 (0.068)	0.214 (0.057)	1.275 (0.076)
36	1.07	0.611 (0.041)	-4.000 (0.230)	-2.285 (0.107)	-0.418 (0.074)	1.713 (0.100)
DS						
2	2.45	1.248 (0.086)	-2.314 (0.106)	-1.125 (0.055)	0.077 (0.045)	1.512 (0.068)
7	1.38	1.060 (0.073)	-2.553 (0.123)	-1.306 (0.061)	-0.149 (0.049)	1.229 (0.064)
12	1.89	1.134 (0.078)	-2.355 (0.108)	-1.221 (0.058)	0.042 (0.048)	1.422 (0.067)
16	1.29	1.148 (0.079)	-2.359 (0.109)	-1.158 (0.056)	0.043 (0.047)	1.431 (0.068)
22	1.38	0.709 (0.048)	-4.267 (0.300)	-2.214 (0.106)	-0.131 (0.068)	1.983 (0.102)
27	1.98	1.116 (0.076)	-2.253 (0.101)	-1.100 (0.056)	0.167 (0.049)	1.610 (0.074)
32	1.04	0.710 (0.047)	-2.784 (0.137)	-1.079 (0.072)	0.476 (0.065)	2.146 (0.105)
37	0.91	1.227 (0.084)	-2.305 (0.106)	-0.993 (0.053)	0.264 (0.046)	1.606 (0.071)
AA						
5	1.64	1.483 (0.105)	-1.133 (0.051)	-0.211 (0.039)	0.708 (0.045)	1.899 (0.081)
8	0.96	1.266 (0.087)	-1.929 (0.081)	-0.969 (0.049)	0.129 (0.044)	1.411 (0.065)
13	1.24	1.667 (0.119)	-1.293 (0.052)	-0.380 (0.038)	0.646 (0.041)	1.745 (0.073)
18	1.02	1.063 (0.073)	-2.686 (0.136)	-1.466 (0.065)	-0.156 (0.050)	1.295 (0.067)
23	1.58	0.740 (0.049)	-3.171 (0.165)	-1.659 (0.081)	0.008 (0.063)	1.635 (0.087)
28	1.72	0.840 (0.057)	-3.769 (0.258)	-1.967 (0.092)	-0.0.43 (0.060)	1.898 (0.094)
34	1.22	0.735 (0.049)	-3.350 (0.184)	-1.645 (0.083)	0.198 (0.065)	2.047 (0.103)
38	1.67	0.966 (0.065)	-2.463 (0.117)	-1.119 (0.062)	0.477 (0.055)	1.954 (0.091)
NJ						
3	1.84	0.878 (0.059)	-1.859 (0.089)	-0.361 (0.054)	0.667 (0.054)	1.840 (0.084)
10	1.55	0.930 (0.063)	-1.628 (0.076)	-0.539 (0.053)	0.697 (0.055)	1.869 (0.084)
14	1.35	1.323 (0.092)	-1.990 (0.086)	-0.933 (0.049)	0.076 (0.042)	1.288 (0.060)
17	1.73	0.982 (0.066)	-1.819 (0.085)	-0.398 (0.052)	0.884 (0.055)	1.952 (0.085)
25	2.60	1.504 (0.106)	-1.895 (0.080)	-0.920 (0.047)	0.327 (0.041)	1.346 (0.058)
30	1.42	1.677 (0.120)	-1.799 (0.073)	-0.946 (0.045)	0.043 (0.037)	0.998 (0.048)
35	1.00	1.061 (0.073)	-2.011 (0.088)	-0.975 (0.055)	0.243 (0.047)	1.240 (0.062)
39	1.96	0.949 (0.065)	-2.074 (0.095)	-0.834 (0.054)	0.141 (0.049)	1.342 (0.069)
NR						
4	1.55	0.585 (0.038)	-3.303 (0.170)	-1.350 (0.087)	0.805 (0.081)	3.137 (0.164)
9	1.91	0.623 (0.041)	-3.490 (0.192)	-1.232 (0.084)	0.729 (0.077)	3.266 (0.177)
19	2.16	0.771 (0.051)	-2.337 (0.110)	-1.056 (0.066)	0.455 (0.062)	2.492 (0.126)
21	1.33	0.660 (0.043)	-3.381 (0.180)	-1.637 (0.088)	0.456 (0.073)	2.923 (0.154)
24	1.51	0.577 (0.038)	-2.167 (0.115)	-0.145 (0.075)	1.813 (0.098)	4.085 (0.241)
29	2.47	1.008 (0.067)	-2.286 (0.111)	-0.773 (0.057)	0.874 (0.058)	2.463 (0.124)
33	2.53	0.873 (0.058)	-2.383 (0.117)	-0.752 (0.062)	1.203 (0.070)	3.095 (0.177)

Note: OB = Observing, DS = Describing, AA = Acting with Awareness, NJ = Non-Judging on Inner Experiences, NR = Non-Reactivity to Inner Experiences, ( ) = Standard Error.

* = $p < .001$

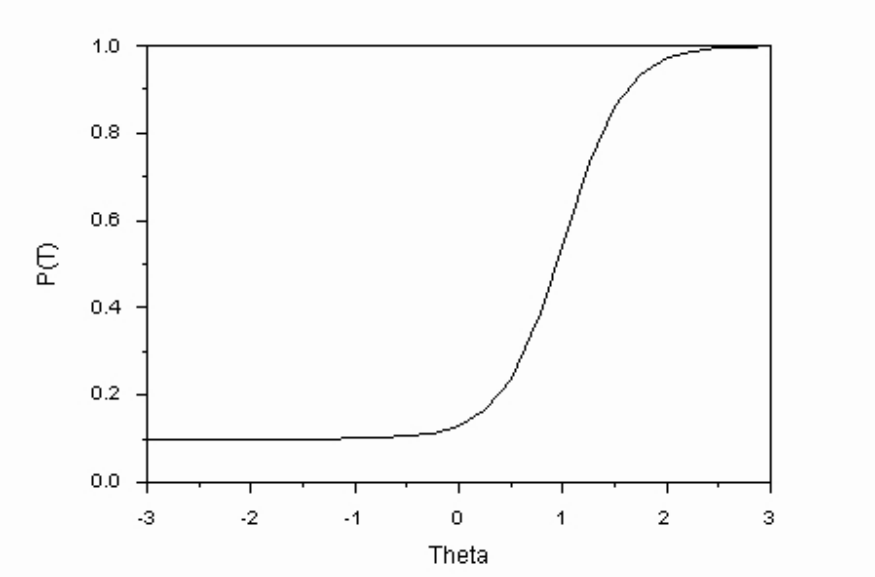

Figure A1. Example of an item characteristic curve where the .50 probability of a correct response is at θ = 1.0. This depicts that individuals at the upper end of the ability scale are more likely to generate correct responses for this item.

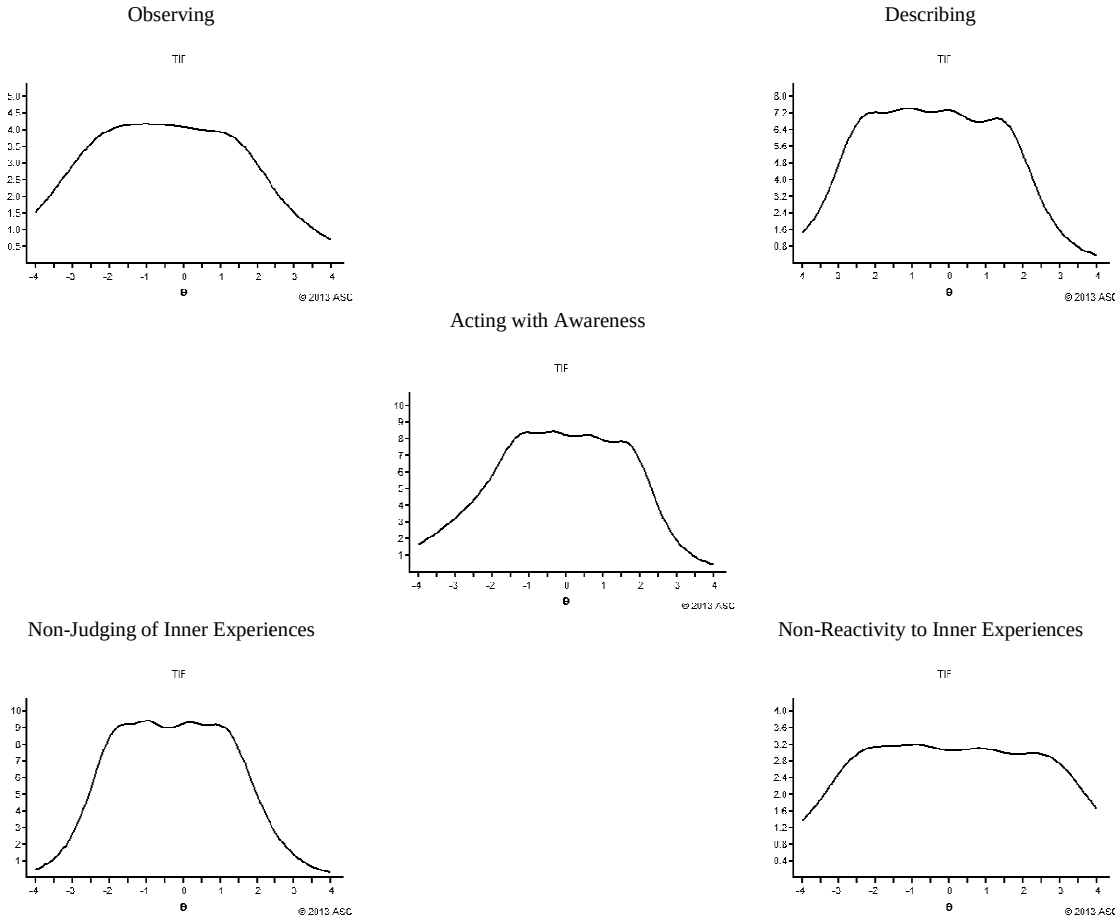

Figure A2. Test information function for each facet of the Five-Facet Mindfulness Questionnaire.

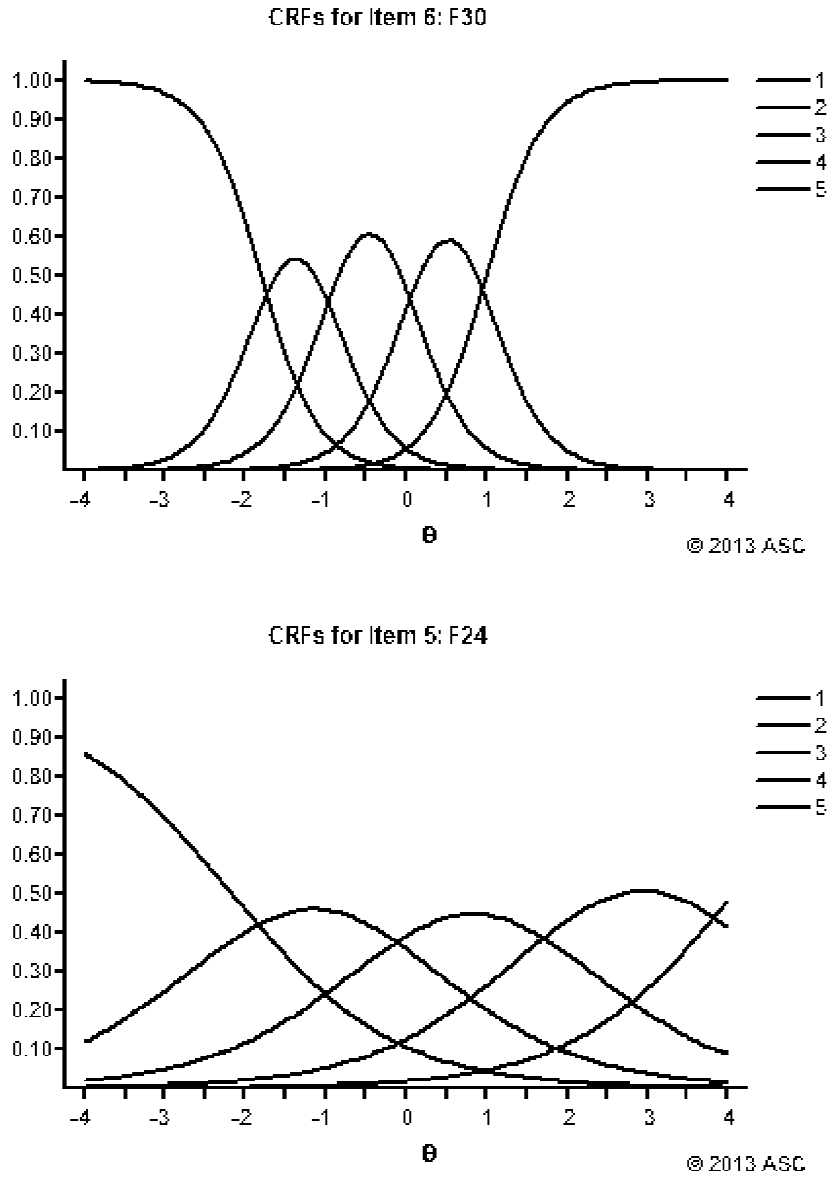

Figure A3. Example categorical response functions (CRF). The CRF of item 30 from the non-judging of inner experiences scale represents the desired pattern for an item with good discrimination. The graphic representing item 24 from the non-reactivity to inner experiences scale demonstrates the pattern present in items with poor discrimination.

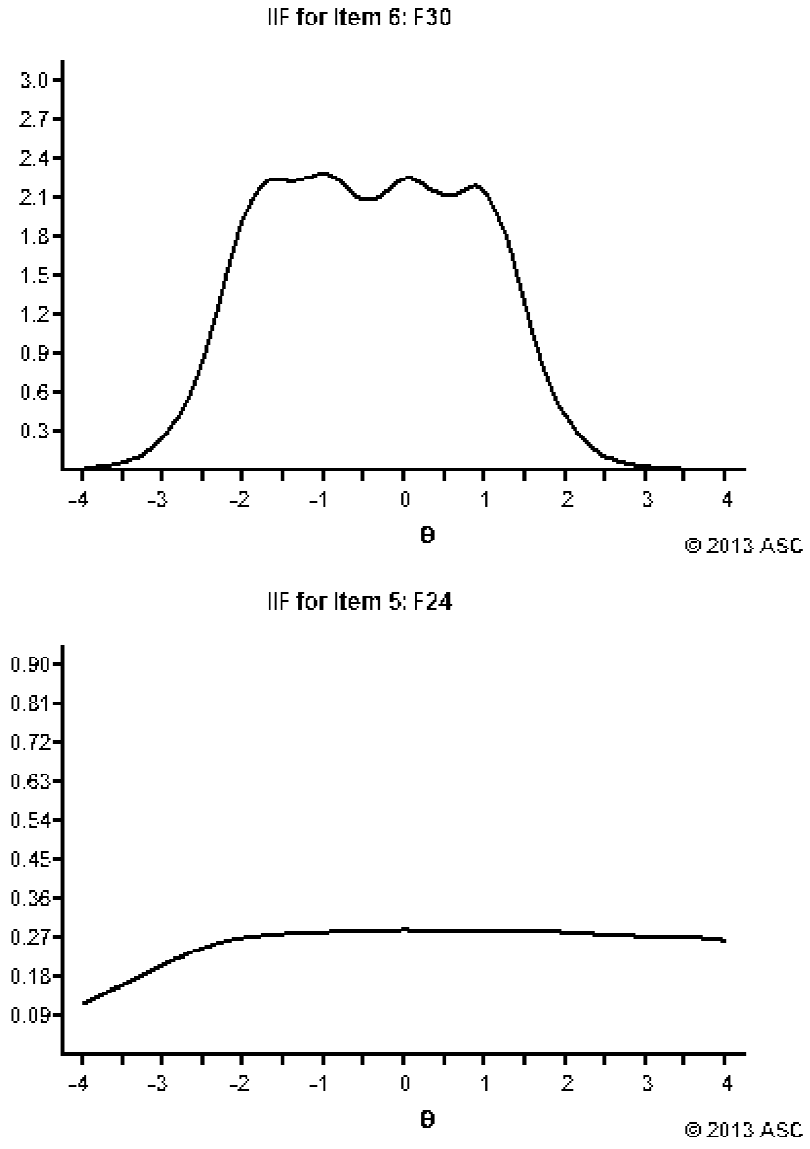

Figure A4. Examples of item information functions (IIF). The IIF for item 30 from the non-judging of inner experiences scale illustrates that information is present at a range of -2.5 to 2.0 on the latent trait. The IIF for item 24 from the non-reactivity to inner experiences scale illustrates the little information is available at any level of the latent trait. Items with more information provide better discrimination.

APPENDIX B: SURVEY ITEMS

Demographics

1. What is your gender?

1) Female
2) Male

2. What is your Age?

3. What is your racial identification?

1) Asian
2) American Indian or Alaskan Native
3) Black or African American
4) Hispanic or Latino
5) White, Non-Hispanic
6) Other

4. Do you currently meditate?

1) Yes
2) No

5. How many days per week do you meditate?

1) 0 to 1
2) 2 to 3
3) 4 to 5
4) 6 to 7

6. Have you ever had any formal training in meditation?

1) Yes
2) No

DASS-42

Please read each statement and circle a number 0, 1, 2 or 3 which indicates how much the statement applied to you *over the past week*. There are no right or wrong answers. Do not spend too much time on any statement. *The rating scale is as follows:*

0 Did not apply to me at all

1 Applied to me to some degree, or some of the time

2 Applied to me to a considerable degree, or a good part of time

3 Applied to me very much, or most of the time

1. I found myself getting upset by quite trivial things.
2. I was aware of dryness of my mouth.
3. I couldn't seem to experience any positive feeling at all.
4. I experienced breathing difficulty (e.g., excessively rapid breathing, breathlessness in the absence of physical exertion).
5. I just couldn't seem to get going.
6. I tended to over-react to situations.
7. I had a feeling of shakiness (e.g., legs going to give way).
8. I found it difficult to relax.
9. I found myself in situations that made me so anxious I was most relieved when they ended.
10. I felt that I had nothing to look forward to.
11. I found myself getting upset rather easily.
12. I felt that I was using a lot of nervous energy.
13. I felt sad and depressed.
14. I found myself getting impatient when I was delayed in any way (e.g., lifts, traffic lights, being kept waiting).
15. I had a feeling of faintness.
16. I felt that I had lost interest in just about everything.
17. I felt I wasn't worth much as a person.

18. I felt that I was rather touchy.
19. I perspired noticeably (e.g., hands sweaty) in the absence of high temperatures or physical exertion.
20. I felt scared without any good reason.
21. I felt that life wasn't worthwhile.
22. I found it hard to wind down.
23. I had difficulty in swallowing.
24. I couldn't seem to get any enjoyment out of the things I did.
25. I was aware of the action of my heart in the absence of physical exertion (e.g., sense of heart rate increase, heart missing a beat).
26. I felt down-hearted and blue.
27. I found that I was very irritable.
28. I felt I was close to panic.
29. I found it hard to calm down after something upset me.
30. I feared that I would be "thrown" by some trivial but unfamiliar task.
31. I was unable to become enthusiastic about anything.
32. I found it difficult to tolerate interruptions to what I was doing.
33. I was in a state of nervous tension.
34. I felt I was pretty worthless.
35. I was intolerant of anything that kept me from getting on with what I was doing.
36. I felt terrified.
37. I could see nothing in the future to be hopeful about.
38. I felt that life was meaningless.
39. I found myself getting agitated.
40. I was worried about situations in which I might panic and make a fool of myself.
41. I experienced trembling (e.g., in the hands).
42. I found it difficult to work up the initiative to do things.

APPENDIX C: RESEARCH MATERIALS

Principal Investigator: Angela Bowman

Study Title: The Psychometric Properties of Mindfulness

Institution: Middle Tennessee State University

The following information is provided to inform you about the research project and your participation in it. Please read this form carefully. You will be given an opportunity to ask questions by using the researcher contact information provided. .

Your participation in this research study is voluntary. You are also free to withdraw from this study at any time by simply ending your participation in the survey.

1. **Purpose of the study:**

   The purpose of this study is to examine the influence of meditation on working memory.

2. **Description of procedures to be followed and approximate duration of the study:**

   You will be seated at a computer and asked to complete a series of questionnaires. Please take your time to read each question carefully and answer truthfully. Participation is expected to last less than 30 minutes.

3. **Expected costs:**

   There will be no cost to you for the data collected for this study.

4. **Description of the discomforts, inconveniences, and/or risks that can be reasonably expected as a result of participation in this study:**

   The risk involved is minimal. It is no more than one would experience in daily life activities. The experiment can be discontinued at any time.

5. **Unforeseeable risks:**

   n/a

6. **Compensation in case of study-related injury:**

   n/a

7. **Anticipated benefits from this study:**

   This study does not provide you with any health care. The study is strictly for research purposes and will have no direct health or medical benefit to you as an individual. The proposed experiments will enable us to address fundamental questions regarding the effectiveness of the questionnaires that you will be answering.

8. **Alternative treatments available:**

   n/a

9. **Compensation for participation:**

   Volunteer participants will not be compensated.

10. **Circumstances under which the Principal Investigator may withdraw you from study participation:**

    n/a.

11. **What happens if you choose to withdraw from study participation:**

    You may decline to join this study or withdraw from this study at any time without prejudice.

12. **Contact Information.** If you should have any questions about your participation in this research study, please feel free to contact Angela Bowman at asb5c@mtmail.mtsu.edu.

13. **Confidentiality.** All efforts, within reason, will be made to keep the personal information in your research record private but total privacy cannot be promised. Your identity will remain confidential. You will be assigned an ID code, and completed forms will be stored in locked files to which only the Principal Investigator will have access. All computer data files pertaining to you will be accessible by subject ID code only. Your IP address will not be logged and no identifying information will be stored. Your information may be shared with MTSU or the government, such as the Middle Tennessee State University Institutional Review Board, Federal Government Office for Human Research Protections, if you or someone else is in danger or if we are required to do so by law.

14. STATEMENT BY PERSON AGREEING TO PARTICIPATE IN THIS STUDY

I have read this informed consent document. I understand each part of the document and I freely and voluntarily choose to participate in this study. I acknowledge that continuing with this survey implies my informed consent to participate.

APPENDIX D: IRB APPROVAL

MIDDLE TENNESSEE STATE UNIVERSITY

October 9, 2012

Angela Bowmen, Dr. Dana Fuller
Department of Psychology
asb5c@mtmtail.mtsu.edu, Dana.Fuller@mtsu.edu

Protocol Title: "Effect of a Single Session of Mindfulness Meditation on Working Memory"
Protocol Number: 13-082

Dear Investigator(s),

The exemption is pursuant to 45 CFR 46.101(b) (2). This is because the research being conducted involves the use of educational tests, survey procedures, interview procedures or public behavior.

You will need to submit an end-of-project report to the Office of Compliance upon completion of your research. Complete research means that you have finished collecting data and you are ready to submit your thesis and/or publish your findings. Should you not finish your research within the three (3) year period, you must submit a Progress Report and request a continuation prior to the expiration date. Please allow time for review and requested revisions. Your study expires on October 9, 2015.

Any change to the protocol must be submitted to the IRB before implementing this change. According to MTSU Policy, a researcher is defined as anyone who works with data or has contact with participants. Anyone meeting this definition needs to be listed on the protocol and needs to provide a certificate of training to the Office of Compliance. **If you add researchers to an approved project, please forward an updated list of researchers and their certificates of training to the Office of Compliance before they begin to work on the project.** Once your research is completed, please send us a copy of the final report questionnaire to the Office of Compliance. This form can be located at www.mtsu.edu/irb on the forms page.

Also, all research materials must be retained by the PI or **faculty advisor (if the PI is a student)** for at least three (3) years after study completion. Should you have any questions or need additional information, please do not hesitate to contact me.

Sincerely,
Andrew W. Jones
Graduate Assistant to:
Emily Born
Compliance Officer
615-494-8918
Emily.Born@mtsu.edu

APPENDIX E: COPYRIGHT APPROVAL