Personality and Individual Differences 49 (2010) 805–810

Contents lists available at ScienceDirect

Personality and Individual Differences

journal homepage: www.elsevier.com/locate/paid

Measuring mindfulness? An Item Response Theory analysis of the Mindful Attention Awareness Scale

Nicholas T. Van Dam [a,*], Mitch Earleywine [a], Ashley Borders [b]

[a] Department of Psychology, University at Albany, SUNY, USA
[b] Department of Psychology, The College of New Jersey, USA

ARTICLE INFO

Article history:
Received 22 March 2010
Received in revised form 15 July 2010
Accepted 16 July 2010
Available online 10 August 2010

Keywords:
Mindfulness
Item Response Theory
MAAS
Construct validity
Meta-consciousness
Psychometrics

ABSTRACT

The Mindful Attention Awareness Scale (MAAS) is one of the most popular measures of mindfulness, exhibiting promising psychometric properties and theoretically consistent relationships to brain activity, mindfulness-based intervention (MBI) outcomes, and mediation of MBI effects. The present study investigated the response patterns and scale properties in a large sample of undergraduate students ($N = 414$) using Item Response Theory analyses. The findings suggest that general statements of "automatic inattentiveness" or "automatic pilot" confer greater statistical information about the underlying latent trait. Evidence of limited abilities to report on mindlessness and of response bias to "mindfulness-absent" items suggests challenges to the construct validity of the MAAS. The current findings, along with preexisting data, suggest that reverse-scoring the scale may be inadequate to represent intentional attention or awareness. Further research is needed to determine which variations, components, and correlates of the numerous operationalizations of mindfulness are theoretically consistent and most salient to positive outcomes, especially in psychopathology.

1. Introduction

Mindfulness has become an increasingly popular construct with diverse clinical and scientific applications (Bishop et al., 2004). Despite efforts at achieving operational definitions and corresponding measurement, disagreement seems the rule rather than the exception (see *Psychological Inquiry*, 2007, Vol. 18, 4). Traditionally, mindfulness involves the active engagement of cognitive-perceptual processes manifest in two broad phases. "The initial phase of mindfulness is the cultivation of sustained bare attention resulting from the practice of non-forgetful attention, followed by... introspective awareness to understand the moment to moment workings of adaptive and maladaptive thoughts and feelings" (Rapgay & Bystrisky, 2009, pp. 153–154). Clinical scientists have attempted to define mindfulness in a way that makes the construct amenable to training and measurement. Bishop and colleagues (2004) provided one of the most integrative and theoretically consistent definitions of the construct.

> The first component involves the self-regulation of attention so that it is maintained on immediate experience, thereby allowing for increased recognition of mental events in the present moment. The second component involves adopting a particular orientation towards one's experiences in the present moment, an orientation that is characterized by curiosity, openness, and acceptance (Bishop et al., 2004, p. 232).

Some have argued that even this elaborate definition fails to represent the true character of mindfulness and has lead to miscomprehension of how mindfulness is developed (Leary & Tate, 2007; Rosch, 2007). Rapgay and Bystrisky (2009) emphasize that mindfulness is an active skill developed by a combination of concentrative and analytical insight-based meditation practices. They also provide an important distinction between attention (a particular cognitive faculty) and awareness (a directable, but broader aspect of consciousness) stating that mindfulness practice entails "...the ability to flexibly apportion...between primary attention to the foreground and secondary awareness to the background..." (p. 155).

There is also debate about therapies related to the construct (e.g., Hofmann & Asmundson, 2008). While many treatments claim a theoretical reliance on mindfulness, the construct has been defined and applied inconsistently (Kabat-Zinn, 2003). Some definitions are based on the Buddhist path towards well-being (e.g., Kabat-Zinn, 1990) while others rely on a reductionist notion of mindfulness (see Hofmann and Asmundson (2008)). Despite theoretical variations, MBIs have been shown to be efficacious treatments for physical and psychological symptoms and conditions (Grossman, Niemann, Schmidt, & Walach, 2004; Hofmann, Sawyer,

* Corresponding author. Address: University at Albany, SUNY, Department of Psychology, Social Sciences 399, 1400 Washington Avenue, Albany, NY 12222, USA. Tel.: +1 (518) 445 5533; fax: +1 (518) 442 4867.
E-mail address: NV122142@albany.edu (N.T. Van Dam).

0191-8869/$ - see front matter © 2010 Elsevier Ltd. All rights reserved.
doi:10.1016/j.paid.2010.07.020

Witt, & Oh, 2010). MBIs often improve health and stress, but change self-reported mindfulness inconsistently (Grossman, 2008). Dramatic variations in operationalizations of mindfulness have led some to question whether scales of "mindfulness" measure the same construct (Rosch, 2007).

Grossman (2008) emphasized several concerns with self-reported mindfulness, including issues of scale construction, potential bias, and item miscomprehension (see also Van Dam, Earleywine, and Danoff-Burg (2009)). Interrelationships among scales purportedly assessing state versus trait mindfulness (e.g., Thompson & Waltz, 2007), as well as behavioral versus self-report mindfulness and predicted outcomes (Frewen, Evans, Maraj, Dozois, & Partridge, 2008) have been inconsistent. Recent evaluation of two of the most popular self-report measures across a Thai and US sample exhibited serious psychometric complications, including no latent mean difference between groups on the Mindful Attention Awareness Scale (MAAS; Brown & Ryan, 2003) despite large differences in meditation and endorsement of Buddhist ideology (Christopher, Charoensuk, Gilbert, Neary, & Pearce, 2009). Further, a recent examination of meditators and non-meditators on the Five Facet Mindfulness Questionnaire (FFMQ; Baer, Smith, Hopkins, Krietemeyer, & Toney, 2006) showed large differential item functioning (Van Dam et al., 2009). While self-report mindfulness scales often have well-established nomothetic span (appropriate correlations with related and unrelated construct measures), they lack construct representationalism (psychological processes underlying responses to a task), an important component of establishing construct validity (Strauss & Smith, 2009).

The MAAS (Brown & Ryan, 2003) is a possible exception to the construct representation problem, with a specific cognitive theory related to scale development. Brown and Ryan (2003) specifically chose items representing *mindlessness* because "...states reflecting less mindlessness are likely more accessible to most individuals, given that mindless states are much more common than mindful states..." (p. 826). The MAAS has also shown theoretically consistent relationships to brain activity (e.g., Creswell, Way, Eisenberger, & Lieberman, 2007), treatment outcome in MBIs (e.g., Michalak, Heidenreich, Meibert, & Schulte, 2008), mediation of targeted MBI outcomes (e.g., Nyklíček & Kuijpers, 2008), and salutary non-targeted benefits resulting from MBIs (Frewen et al., 2008). The MAAS has a strongly supported unidimensional factor structure and good nomothetic span (e.g., Brown & Ryan, 2003; MacKillop & Anderson, 2007), making it a seemingly good candidate to represent mindfulness.

Unfortunately, the assumption regarding the accessibility of mindless states is challenged by empirical investigation in cognitive neuroscience. Recent studies in meta-awareness and attention suggest that mind wandering (typically a mindless state) is associated with a lack of meta-awareness (awareness that one is not aware). Further, attention is decoupled from task engagement during a mind-wandering episode (Smallwood, McSpadden, & Schooler, 2007). These findings suggest that one's ability to accurately report about mindless states may be limited without specific training. Lack of meta-awareness regarding mindless states (or mindfulness-absent states) suggests that responses on the MAAS are likely not the result of the proposed cognitive process. Further, a recent examination of mindfulness-absent items on the FFMQ suggests a general response bias to reject items suggesting higher prevalence of *mindlessness* (Van Dam et al., 2009), revealing construct-inconsistent response processes.

1.1. Applying Item Response Theory

Given direct challenges to construct validity and the underlying response processes, Item Response Theory (IRT) is well suited to provide information potentially not available via Classical Test Theory analyses (see de Ayala, 2009; Embretson & Reise, 2000). The psychometric properties obtained from IRT analyses are theoretically sample invariant, a potentially important consideration in measuring mindfulness given known sample differences (e.g., Christopher et al., 2009; Van Dam et al., 2009). One important psychometric property is the item information estimate. An item's information provides an index of how useful it is in discriminating between participants at specified trait levels. Another important aspect of IRT is the detailed prediction of responses based on estimated latent trait level. The relationship of an individual's estimated trait level (θ) and the probability of choosing a given response is exemplified on an item-by-item basis by a category response curve (CRC) (see Fig. 1).

The MAAS has a polytomous response format (e.g., a six-point rating scale), with graded response options. A theoretically appropriate IRT model to explore this item response format is the Graded Response Model (GRM; see de Ayala (2009), Embretson and Reise (2000), Ostini and Nering (2006), Samejima (1969) and Samejima (1996)). The GRM entails cumulative response distributions for each item, computing thresholds between each response option and all options ordinally higher (Samejima, 1969). Each CRC predicts the probability of all response options, anchored at a given trait level by b_i (the threshold parameter). The threshold parameter, b_i, is that point along the trait continuum where the probability of selecting a given response option is 50%. The GRM also permits estimation of an item discrimination parameter (a_i) for each item. The item discrimination parameter provides an estimate of how well the item differentiates between individuals of varying trait levels. Values of a_i from 0.01 to 0.24 are considered very low, 0.25–0.63 low, 0.65–1.34 moderate, 1.35–1.69 high, and >1.7, very high (Baker, 2001).

1.2. Goodness of fit

Empirical and simulation studies exploring model fit in IRT have provided no evidence of a "gold standard", however several statistical indices have been proposed (Drasgow, Levine, & Williams, 1985; Hambleton, Swaminathan, & Rogers, 1991; Ostini & Nering, 2006). Hambleton and colleagues (1991) suggest an exploration of model assumptions followed by cautious use of statistical tests. All basic IRT models assume data unidimensionality, which can be examined with confirmatory factor analysis based on previously specified scale factor structure. The GRM also assumes invariant trait (θ) parameter estimates as well as invariant item parameter estimates (a_i, b_i). A cursory examination of these model features can be conducted by comparing trait estimates from even and odd items as well as parameter estimates from male and female participants, respectively (see Hambleton et al. (1991, p. 64)). If model assumptions are not violated, the relationship between model predictions and actual data can be explored using residual-based measures (e.g., Drasgow et al., 1985; Embretson & Reise, 2000). Simulations and empirical data suggest that the Z_L statistic developed by Drasgow and colleagues (1985) is a particularly robust estimator of person-fit and model-fit (Reise, 1990).

1.3. Current study

Data unidimensionality was assessed by confirmatory factor analysis (CFA) in LISREL v 8.8 (Jöreskog & Sörbom, 1993). Cutoff criteria were established for reasonable and good model fit based on recommendations (Brown, 2006; Hu & Bentler, 1999; Marsh, Hau, & Wen, 2004). Standardized root mean residual (SRMR) <.08 was considered good, <.10 was considered reasonable; comparative fit index (CFI) >.95 was considered good, >.90 was considered reasonable; root mean square error of approximation <.06 was considered good, <.08 was considered reasonable. To explore the latent trait

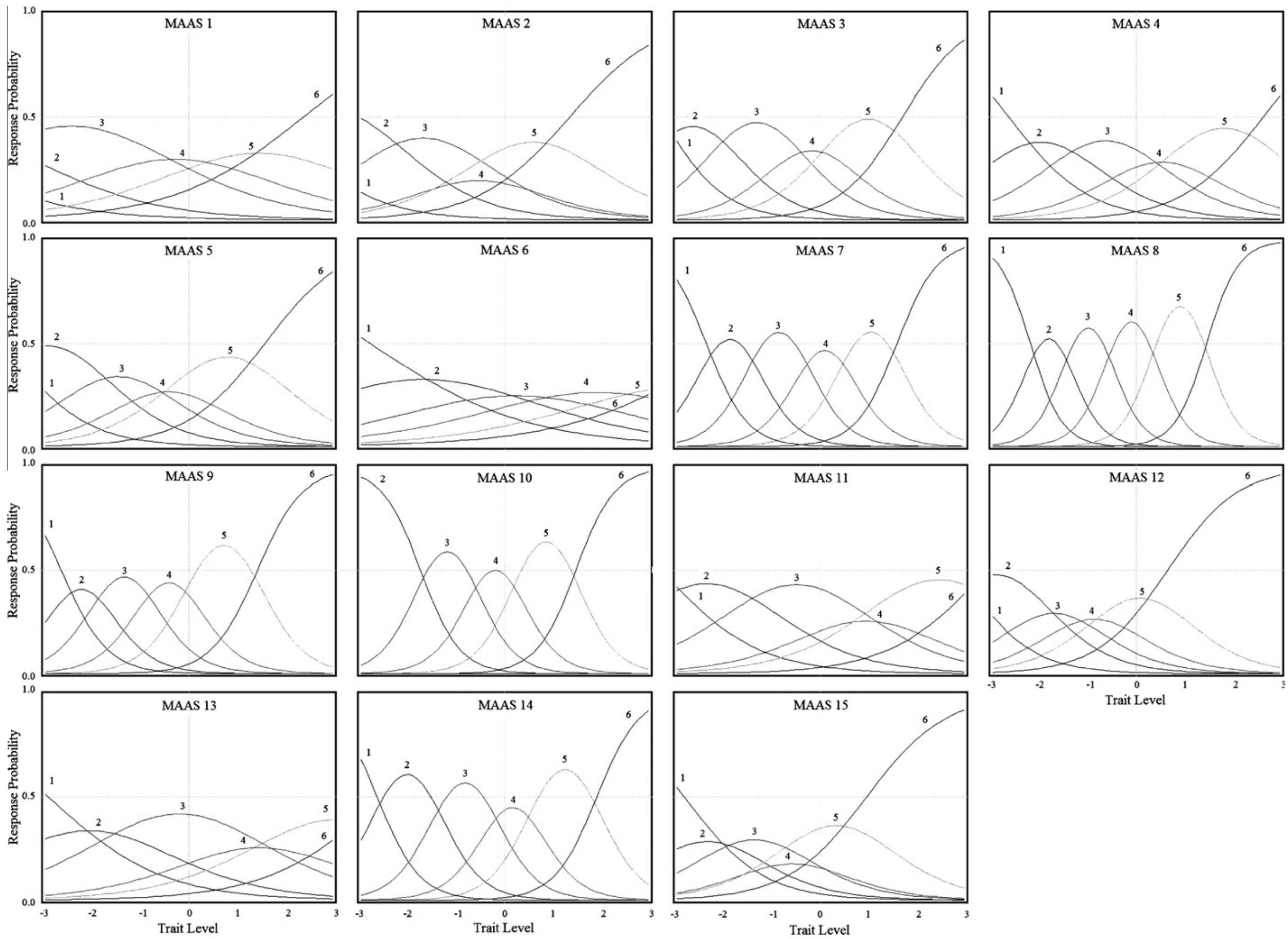

Fig. 1. Category response curves for all 15 items of the Mindful Attention Awareness Scale. Category response curves (CRC) represent the probability of a participant with trait level θ, endorsing a particular response option. The b_i and a_i parameters, estimated by the Graded Response Model (GRM) dictate the specific location and shape of the CRCs. The threshold parameter b_i corresponds to the point of intersection between the CRC for each response option and the CRC for the next ordinally higher response option. Note the relatively high CRC overlap for all items except 7, 8, 9, 10, and 14.

underlying responses to the MAAS, item parameters and individual trait estimates were computed using MULTILOG v 7.03 (Thissen, 1991). All other statistics were computed using SPSS 17.0. Goodness-of-fit was assessed via the Z_L statistic (Drasgow et al., 1985).

2. Methods

2.1. Participants and procedure

Four-hundred-fourteen (414) undergraduate students at a major university in Los Angeles completed several measures given to students enrolled in psychology courses. The local institutional review board approved all procedures. Participants' mean age was 20.1 (SD = 2.54) and two-thirds were female. Just over 50% of participants self-identified as Caucasian, 19% as Asian, 9% as Hispanic, 7% as African-American, and 13% as "other."

2.2. Measures

2.2.1. Mindfulness

The Mindful Attention Awareness Scale is a 15-item questionnaire that according to Brown and Ryan (2003) measures dispositional mindfulness (e.g., "I rush through activities without being really attentive to them" and "I drive places on 'automatic pilot' and then wonder why I went there"). Participants reported how often they believed they had experiences referenced by each item on a 6-point Likert scale from "almost always" to "almost never". Higher averaged total scores purportedly reflect higher mindfulness. In this sample, Cronbach's alpha for the measure was .88.

3. Results

Data were normally distributed (skewness and kurtosis values <1 for all individual items), with overall scale skewness = −.084 (SE = .120), and kurtosis = −.273 (SE = .239). There was one univariate outlier, MAAS total Z score > ±3.29. There were 13 multivariate outliers, Mahalanobis distances above the cut-score ($p < .001$). All 14 outliers were excluded from further analyses, leaving 400 cases. There were 10 incidences of missing data (0.17% of total responses). Proving that data is missing completely at random is controversial. Accordingly, missing values were imputed using maximum-likelihood estimation with subsequent rounding to a whole integer to better mimic actual responses.

3.1. Goodness of fit

CFA with maximum-likelihood estimation supported data unidimensionality. The fit indices suggest a good fit between the single factor model and observed data, $\chi^2_{(90)} = 237.38$, standardized root mean square residual (SRMR) = .046, comparative fit index (CFI) = 0.971, root mean square error of approximation (RMSEA) = 0.065. Two of three fit indices met criteria for good fit (SRMR and CFI) while the third (RMSEA) met reasonable fit criteria. The 90% confidence interval for the RMSEA ranged from 0.056 to 0.075, further suggesting good fit (Brown, 2006).

Trait estimates were computed separately for even and odd items to explore the IRT assumption of trait invariance (Hambleton et al., 1991). The correlation between the independent trait estimates suggests no violation of trait invariance, $r = .833$, $p < .01$. Item parameter estimates were computed independently for males and females (Hambleton et al., 1991). The correlation between the a_i parameter by gender suggests no violation of parameter invariance, $r = .933$, $p < .01$.

The Z_L statistic (Drasgow et al., 1985) was computed to assess model prediction of actual scores. Across all responses for each item, only 156 of 6000 (2.6%) standardized residuals (Z_L) were greater than ±1.92 (cut-off value for $p < .05$). The overall average absolute Z_L was 0.65, with a variance of 0.51. The largest average absolute Z_L for an individual was 1.68 ($p = .09$). These numbers fall within the expectations for high overall goodness-of-fit and person goodness-of-fit criteria (Reise, 1990).

3.2. Estimated item parameters

All item parameters are displayed in Table 1. Item discrimination parameters (a_i) span from 0.67 to 2.87 (logistic metric) with a mean value of 1.51 (MDN = 1.3), suggesting moderate to very high discrimination (Baker, 2001). Note that a large a_i parameter value cannot be interpreted without considering the latent trait range within which the item discriminates. Threshold parameters (b_i) were unevenly distributed across the trait range, with many items exhibiting strong positive skew, indicating that most individuals are unlikely to endorse lower response options. Six items were distributed evenly across the trait range (items 4, 7, 8, 9, 10, and 14; see Table 1), suggesting that only these items differentiate individuals at low through high trait levels.

Table 2 displays the amount of information that each item yields on its own and in proportion to the total scale. Simultaneously examining Table 2 and Fig. 1, one can see that low information corresponds to broad, undifferentiating CRCs (e.g., see item 6). Items 7, 8, 9, 10, and 14 contribute approximately 2/3 of the overall scale information (see Fig. 2), suggesting that other items contribute relatively little to scale utility. Item 3, which provides the most information of all other items ($I = 18.32$), yields 40% less information than the least informative of the five items described above, item 9 ($I = 30.67$). Items 7, 8, 9, 10, and 14 had corrected item-total correlations ranging from $r = .64$ to $r = .74$, with an average $r = .69$. Corrected item-total correlations for all other items ranged from $r = .34$ to $r = .57$, with an average $r = .46$, suggesting that these items have less relation to the latent trait that the scale represents. Internal consistency for items 7, 8, 9, 10, and 14 was only .003 Cronbach's α units less than for the total scale. Clearly, these items strongly reflect the scale trait.

Table 1
Item Response Theory parameter estimates for the Mindful Attention Awareness Scale.

	Item parameter estimates[a]					
Item	a	b_1	b_2	b_3	b_4	b_5
1	0.86	−5.21	−3.22	−0.93	0.38	2.11
2	0.92	−5.38	−2.69	−0.92	−0.08	1.59
3	1.47	−3.45	−2.03	−0.63	0.30	1.72
4	1.14	−2.68	−1.39	−0.07	0.90	2.79
5	1.30	−3.52	−1.83	−0.74	0.14	1.65
6	0.67	−2.29	−0.47	1.11	2.70	4.27
7	2.36	−2.29	−1.31	−0.31	0.58	1.70
8	2.87	−2.20	−1.39	−0.50	0.43	1.58
9	2.00	−2.51	−1.66	−0.77	0.19	1.56
10	2.45	−2.62	−1.49	−0.53	0.33	1.56
11	1.00	−3.14	−1.39	0.35	1.48	3.35
12	1.38	−3.61	−2.23	−1.15	−0.36	0.71
13	0.89	−2.78	−1.04	0.69	2.05	3.72
14	2.17	−2.86	−1.47	−0.28	0.63	2.00
15	1.12	−3.21	−2.17	−1.06	−0.43	0.88

[a] Each of the b parameters corresponds to a probability = 0.5 of choosing the response that is +1 from the subscript. The a parameter is the slope at the location of all b parameters and corresponds to the item's ability to discriminate between individuals of different trait levels.

4. Discussion

This study used IRT analyses to examine responses on the Mindful Attention Awareness Scale (MAAS), one of the most popular self-report measures of mindfulness. The majority of items exhibited broad and indiscriminate CRCs (see Fig. 1), suggesting an inability of response options to differentiate between trait levels. A small subset of items exhibited a reasonable span across the trait range; all exhibiting very high discrimination parameters (Baker, 2001). These five items (7, 8, 9, 10, and 14) appear to provide the majority of the information (ability to discriminate between individuals of varying levels of the trait for the MAAS). Adding the other 10 items provides more overall information, but does not substantially alter curve shape (Fig. 2) or scale reliability and actually decreases item-total correlations.

Examining the anchor points of the response thresholds (see Table 1) and the overall distribution of information (see Fig. 2) reveals that information below $\theta = -2$ and above $\theta = 1.5$ is poorly represented relative to the rest of the trait range. Accordingly, even the most psychometrically solid items of the MAAS seem incapable of differentiating high from very high as well as low from very low levels of the latent trait. A comparable inability to discern between trait levels appeared in the lack of mean latent variable difference on the MAAS between a Thai and American sample, where differences in meditation experience and Buddhist values were striking (Christopher et al., 2009).

The content of items 7, 8, 9, 10, and 14 suggests that item information may be a function of statement generality. In comparison to the other items, these five items are not anchored to specific experiences or behaviors (e.g., interpersonal interaction, driving, eating). The fact that more general items exhibit better discrimination between varying trait levels may partially be due to the difficulty associated with self-report of mindless states (e.g., Smallwood et al., 2007). Since it is inherently difficult to self-report about states of which one is unaware, it may be easier to respond to more general items.

Generally consistent with the developers' original proposition (Brown & Ryan, 2003), the above five items seem to relate to "automatic inattentiveness" or "automatic pilot" (see Kabat-Zinn, 1990). However, a reversal of total scale score seems an unlikely means of measuring the opposite (mindfulness) of what the items represent, in contrast to Brown and Ryan's (2003) proposal. If the latent trait is "automatic inattentiveness", it could suggest something about attention regulation, the first component of Bishop and colleagues' (2004) operationalization of mindfulness. However, the small subset of items neither states nor implies anything about one's intention or state of mind, an important component of the practices that enhance mindfulness (Rapgay & Bystrisky, 2009) and part of Bishop et al.'s (2004) operationalization. Another important limitation is that the MAAS does not differentiate attention and awareness, interacting, but arguably distinct components of

Table 2
MAAS item information summed across trait estimates.

Item	Information	% Total information
1. I could be experiencing some emotion and not be conscious of it until some time later	6.84	2.20
2. I break or spill things because of carelessness, not paying attention, or thinking of something else	7.52	2.41
3. I find it difficult to stay focused on what's happening in the present	18.32	5.88
4. I tend to walk quickly to get where I'm going without paying attention to what I experience along the way	11.97	3.84
5. I tend not to notice feelings of physical tension or discomfort until they really grab my attention	14.40	4.63
6. I forget a person's name almost as soon as I've been told it for the first time	4.26	1.37
7. It seems I am "running on automatic," without much awareness of what I'm doing	**41.03**	**13.18**
8. I rush through activities without being really attentive to them	**54.80**	**17.60**
9. I get so focused on the goal I want to achieve that I lose touch with what I'm doing right now to get there	**30.67**	**9.85**
10. I do jobs or tasks automatically, without being aware of what I'm doing	**43.35**	**13.92**
11. I find myself listening to someone with one ear, doing something else at the same time	9.38	3.01
12. I drive places on 'automatic pilot' and then wonder why I went there	14.41	4.63
13. I find myself preoccupied with the future or the past	7.55	2.42
14. I find myself doing things without paying attention	**36.64**	**11.77**
15. I snack without being aware that I'm eating.	10.23	3.29

Note that together the items in bold represent approximately 2/3 of the total information of the scale.

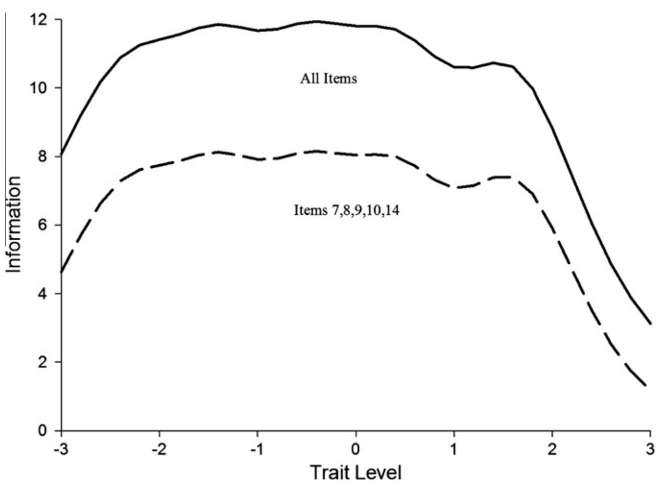

Fig. 2. Total information for the Mindful Attention Awareness Scale across trait estimates. Information is determined for each item at each trait level at each response threshold, where there are $k-1$ response thresholds (k = total number of response options). Total information is the sum of information across all trait levels and all response options for each item. In this figure, total scale information is represented by the solid line and information for a select subset of items (7, 8, 9, 10, 14) is represented by the dotted line. Note that information is relatively consistent between trait levels −2 and 1.5, but declines sharply below $\theta = 2$ and above $\theta = 1.5$.

consciousness (e.g., Rapgay & Bystrisky, 2009; Smallwood et al., 2007).

Rosch (2007) has suggested that something like level of psychopathology may underlie responses on mindfulness scales. Indeed, the MAAS has consistently exhibited significant negative correlations with broad psychological constructs commonly related to psychopathology (e.g., neuroticism – Brown & Ryan, 2003; Thompson & Waltz, 2007, negative intrusive thoughts and/or rumination – Frewen et al., 2008, and negative affect – Brown & Ryan, 2003; Nykliček & Kuijpers, 2008). That the MAAS may reflect a broad psychological construct is supported by the general nature of the five most statistically informative statements in the current study.

Despite these results, findings that measures of mindfulness relate to MBI outcomes (Michalak et al., 2008) and appropriate neural activity (e.g., Creswell et al., 2007) should not be summarily dismissed. As Strauss and Smith (2009) note, construct validity is difficult to establish when there is no basis for comparison. Efforts at defining and measuring mindfulness have been laudable. Nonetheless, if the goal is to capitalize on the benefits of mindfulness, it behooves researchers to use a definition of mindfulness that has theory-based construct validity. Finding and operationalizing this definition may require more input from those familiar with both traditional Buddhist conceptualizations of the notion as well as an understanding of its modern psychological implications (e.g., Christopher et al., 2009; Grossman, 2008; Kabat-Zinn, 2003; Leary & Tate, 2007; Rapgay & Bystrisky, 2009; Rosch, 2007). Improved or alternative scales might not only better represent the construct, but could also uncover other benefits of mindfulness practice and larger effect sizes. This approach would also permit improvements in MBIs, targeting the most effective and efficacious components and correlates of mindfulness. Difficulties inherent to re-representing consciousness may require exploration of other constructs closely related to and encouraged by MBIs as well (Grossman, 2008; Van Dam et al., 2009). Limiting assessment solely to a Westernized version of a complex Buddhist concept may be shortsighted in consideration of a fundamentally different way of *being* and commitment to a path of well-being.

Acknowledgements

The authors thank Leo Stam for programming assistance, Lobsang Rapgay for helpful discussion of the construct, and Gwyneth Boodoo, Patrick Jones, and two anonymous reviewers for comments on earlier drafts.

References

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

Baker, F. (2001). *The basics of item response theory. ERIC clearinghouse on assessment and evaluation*. MD: University of Maryland College Park.

Bishop, S. B., Lau, M., Shapiro, S., Carlson, L., Anderson, N. D., Carmody, J., et al. (2004). Mindfulness: A proposed operational definition. *Clinical Psychology: Science and Practice, 11*, 230–241.

Brown, T. A. (2006). *Confirmatory factor analysis for applied research*. New York: Guilford.

Brown, K. W., & Ryan, R. M. (2003). The benefits of being present: Mindfulness and its role in psychological well-being. *Journal of Personality and Social Psychology, 84*(2), 822–848.

Christopher, M. S., Charoensuk, S., Gilbert, B. D., Neary, T. J., & Pearce, K. L. (2009). Mindfulness in Thailand and the United States: A case of apples versus oranges? *Journal of Clinical Psychology, 65*, 590–612.

Creswell, J. D., Way, B. M., Eisenberger, N. I., & Lieberman, M. D. (2007). Neural correlates of dispositional mindfulness during affect labeling. *Psychosomatic Medicine, 69*, 560–565.

de Ayala, R. J. (2009). *The theory and practice of item response theory*. New York: Guilford Press.

Drasgow, F., Levine, M. V., & Williams, E. A. (1985). Appropriateness measurement with polychotomous item response models and standardized indices. *British Journal of Mathematical and Statistical Psychology, 38*, 67–86.

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

Frewen, P. A., Evans, E. M., Maraj, N., Dozois, D. J. A., & Partridge, K. (2008). Letting go: Mindfulness and negative automatic thinking. *Cognitive Therapy and Research, 32*, 758–774.
Grossman, P. (2008). On measuring mindfulness in psychosomatic and psychological research. *Journal of Psychosomatic Research, 64*, 405–408.
Grossman, P., Niemann, L., Schmidt, S., & Walach, H. (2004). Mindfulness-based stress reduction and health benefits: A meta-analysis. *Journal of Psychosomatic Medicine, 57*, 35–43.
Hambleton, R. K., Swaminathan, H., & Rogers, H. J. (1991). *Fundamentals of item response theory*. Newbury Park, CA: Sage Publications, Inc..
Hofmann, S. G., & Asmundson, G. J. G. (2008). Acceptance and mindfulness-based therapy: New wave or old hat? *Clinical Psychology Review, 28*, 1–16.
Hofmann, S. G., Sawyer, A. T., Witt, A. A., & Oh, D. (2010). The effect of mindfulness-based therapy on anxiety and depression: A meta-analytic review. *Journal of Consulting and Clinical Psychology, 78*, 169–183.
Hu, L., & Bentler, P. M. (1999). Cutoff criteria for fit indices in covariance structure analysis: Conventional criteria versus new alternatives. *Structural Equation Modeling, 6*, 1–55.
Jöreskog, K., & Sörbom, D. (1993). *LISREL (version 8.80) [computer software]*. Lincolnwood, IL: Scientific Software International.
Kabat-Zinn, J. (1990). *Full catastrophe living: Using the wisdom of your body and mind to face stress, pain, and illness*. New York: Dell Publishing.
Kabat-Zinn, J. (2003). Mindfulness-based interventions in context: Past, present, and future. *Clinical Psychology: Science and Practice, 10*, 144–156.
Leary, M. R., & Tate, E. B. (2007). The multi-faceted nature of mindfulness. *Psychological Inquiry, 18*(4), 251–255.
MacKillop, J., & Anderson, E. J. (2007). Further psychometric validation of the mindful attention awareness scale (MAAS). *Journal of Psychopathology and Behavioral Assessment, 29*(4), 289–293.
Marsh, H. W., Hau, K.-T., & Wen, Z. (2004). In search of golden rules: Comment on hypothesis- testing approaches to setting cutoff values for fit indices and dangers to overgeneralizing Hu and Bentler's (1999) findings. *Structural Equation Modeling, 11*, 320–341.
Michalak, J., Heidenreich, T., Meibert, P., & Schulte, D. (2008). Mindfulness predicts relapse/recurrence in major depressive disorder after mindfulness-based cognitive therapy. *Journal of Nervous and Mental Disease, 196*(8), 630–633.
Nyklíček, I., & Kuijpers, K. F. (2008). Effects of mindfulness-based stress reduction intervention on psychological well-being and quality of life: Is increased mindfulness indeed the mechanism? *Annals of Behavioral Medicine, 35*, 331–340.
Ostini, R., & Nering, M. L. (2006). *Polytomous item response theory models*. Thousand Oaks, CA: Sage Publications, Inc..
Rapgay, L., & Bystrisky, A. (2009). Classical mindfulness: An introduction to its theory and practice for clinical application. *Annals of the New York Academy of Sciences, 1172*, 148–162.
Reise, S. P. (1990). A comparison of item- and person-fit methods of assessing model-data fit in IRT. *Applied Psychological Measurement, 14*, 127–137.
Rosch, E. (2007). More than mindfulness: When you have a tiger by the tail, let it eat you. *Psychological Inquiry, 18*(4), 258–264.
Samejima, F. (1969). Estimation of latent ability using a response pattern of graded scores. *Psychometric Monograph Supplement*, *17*(4, Pt 2).
Samejima, F. (1996). Evaluation of mathematical models for ordered polychotomous responses. *Behaviormetrika, 1*, 17–35.
Smallwood, J., McSpadden, M., & Schooler, J. W. (2007). The lights are on but no one's home: Meta-awareness and the decoupling of attention when the mind wanders. *Psychonomic Bulletin and Review, 14*, 527–533.
Strauss, M. E., & Smith, G. T. (2009). Construct validity: Advances in theory and methodology. *Annual Review of Clinical Psychology, 5*, 1–25.
Thissen, D. (1991). *MULTILOG (version 7.03) [computer software]*. Lincolnwood, IL: Scientific Software Incorporated.
Thompson, B. L., & Waltz, J. (2007). Everyday mindfulness and mindfulness meditation: Overlapping constructs or not? *Personality and Individual Differences, 43*, 1875–1885.
Van Dam, N. T., Earleywine, M., & Danoff-Burg, S. (2009). Differential item function across meditators and non-meditators on the Five Facet Mindfulness Questionnaire. *Personality and Individual Differences, 47*, 516–521.