Psychometric Analysis of the Multidimensional State Boredom Scale and its Condensed Versions

Abstract

The Multidimensional State Boredom Scale is a promising new self-report measure of state boredom. Two condensed versions of the scale have also been introduced. This study helped to explore the psychometric qualities of these scales, using a large sample of Australian adults (N = 1716), as well as two smaller samples (N = 199 and N = 422). Data analyses indicated strong convergent validity and very high internal consistency for the scales. Test-retest reliability over a 6-8 day period was moderately high. Confirmatory factor analyses of the Multidimensional State Boredom Scale authors' suggested factor structure indicated good fit for this model. However, some of the data analyses raise questions as to whether the scale includes meaningful sub factors. Overall, the MSBS (and Short Form) is recommended for researchers who wish to assess state boredom .

Keywords: Boredom, State, Confirmatory Factor Analysis, Psychometric, Test-retest Reliability, Validity

Psychometric Analysis of the Multidimensional State Boredom Scale and its Condensed Versions

Research has indicated that boredom and/or boredom proneness are a possible contributor to various behavioural problems and health-related issues. These include drug use (Wiesner, Windle, & Freeman, 2005; Woodall, 2012), depression (Ahmed, 1990; Newell, Harries, & Ayers, 2011; Vodanovich, Verner, & Gilbride, 1991), impulse control problems (Blaszczynski, McConaghy, & Frankova, 1990; Wiesbeck et al., 1996), violence (Chaplin, McGeorge, & Lelliott, 2006; Meehan, McIntosh, & Bergen, 2006) , and counterproductive work behaviours (Bruursema, Kessler, & Spector, 2011). Given the growing evidence that boredom is a problematic and widespread phenomenon, additional research is warranted on the nature of boredom, its prevalence, and its impacts on people. To assist in these studies, there are now a range of self-report boredom measures for researchers to employ (see Vodanovich and Watt, 2015, for a review of these measures).

Distinguishing between State and Trait boredom

Traditionally, most measures of boredom have been 'trait' measures of boredom, assessing the personality traits associated with boredom experience (e.g., Farmer & Sundberg, 1986; Zuckerman, 1994). A recent development in the field of boredom measurement is the introduction of general use 'state' measures of boredom, including the Multidimensional State Boredom Scale (Fahlman, Mercer-Lynn, Flora, & Eastwood, 2013) and the State Boredom Measure (Todman, 2013). The conceptual separation between the general tendency to experience boredom (trait) versus a person's situational experience of boredom (state) has been raised by various researchers (e.g. Fahlman et al., 2013; Todman, 2013; Vodanovich, 2003). On the basis of this recent discussion, trait boredom could be described as a consistent internal personality characteristic - a person's tendency to experience boredom, ignoring the influence of any situations they have experienced.

Conversely, state boredom may be defined as a person's subjective experience of boredom during a specific timeframe. To further delineate state boredom, the specific time of interest may be the present (present-state boredom), as assessed by the Multidimensional State Boredom Scale (Fahlman et al., 2013), or in the past (past-state boredom), as assessed by the State Boredom Measure (Todman, 2013). That is, 'present-state boredom' is one's experience/degree of boredom at the time of boredom assessment. Conversely, 'past-state boredom' represents specific boredom experiences over a specific time period in the past. For example, if a person is assessed one morning and is bored during the assessment, they should report high present-state boredom. If they are assessed again that evening when they are not feeling bored, then they should report no present-state boredom. However, they should report at least some level of past-state boredom, assuming the past-state boredom survey includes that morning as a substantial fraction of the assessed timeframe.

A substantial benefit of state boredom measures is that they allow for temporal hypotheses to be tested, which could not be achieved with trait measures. For example, although previous research has linked boredom to violent behaviour (Bensley, Nelson, Kaufman, Silverstein, & Shields, 1995; Loubser, Chaplin, & Quirk, 2009), state measures of boredom could be used to identify whether boredom actually precedes and causes violent behaviour.

About the Multidimensional State Boredom Scale

Given the importance of state measures of boredom, it is essential that these measures receive adequate psychometric analysis using a range of samples. This article focuses on providing psychometric analysis of the Multidimensional State Boredom Scale (MSBS; Fahlman et al., 2013), as well as two condensed versions: The MSBS Short Form (Hunter, Dyer, Cribbie, & Eastwood, 2015), and MSBS 15 item version (MSBS-15; Baratta & Spence, 2015).

The full MSBS (Fahlman et al., 2013) assesses five dimensions of the experience of boredom: Disengagement, High Arousal, Inattention, Low Arousal, and Time Perception. The survey consists of 29 items that are split between these five dimensions. For example, one item in the Disengagement subscale states “I wish I was doing something more exciting”. Each item has a likert-scale response format, ranging from 1 (strongly disagree) to 7 (strongly agree). Higher scores on the MSBS therefore reflect greater boredom in the individual. Importantly, the MSBS measures boredom at the time of measurement (i.e., present state boredom). As such, the context of measurement (e.g., what participants were doing before completing the survey) should have a large influence on a participant’s MSBS scores.

The MSBS-15 (Baratta & Spence, 2015) contains 15 of the items from the original MSBS, including items from each of the original five subscales. The MSBS is scored similarly to the original scale. In creating the MSBS-15, Baratta and Spence used Item Response Theory analyses to investigate which of the original MSBS items had the best psychometric qualities. In particular, the items’ discrimination parameters and item information functions were focused on when retaining items for this scale. The Short Form of the MSBS (Hunter et al., 2015) includes eight of the original items and is intended to differentiate ‘bored’ from ‘not bored’ participants. The Short Form is also scored similarly to the original scale. However, the Short Form contains mostly ‘disengagement’ items, and is not intended to capture the full five dimensions of boredom assessed by the full scale MSBS.

The MSBS was initially developed through testing with Canadian undergraduate samples (Fahlman et al., 2013). The MSBS items were initially developed on the basis of written descriptions of boredom, and incorporated existing theoretical discussion of the definition of boredom. Items were then tested multiple times using several large sample studies, and a series of statistical analyses (e.g., EFA, analysis of skew) helped to select the final 29 items. Construct validity of the MSBS was assessed by Fahlman et al. through experimental

manipulation of boredom, and via comparisons with various related scales/constructs (e.g.,, depression). Fahlman and colleagues' testing showed the scale has quite good internal consistency, with Cronbach's α ranging from .80 to .97 for the full scale and .75 to .93 for subscales.e.g.,

Other research on the psychometric qualities of the MSBS has also provided good support for the scale (e.g., Baratta & Spence, 2015; Hunter et al., 2015). In particular, using structural equation modeling, Baratta and Spence found good support for multidimensional structures for the MSBS-15, in contrast to weaker support for a unidimensional structure. Hunter et al. found evidence of sensitivity and specificity of the MSBS in identifying people who went through a boredom-inducing experimental manipulation. Additionally, Hunter et al. demonstrated gender invariance for all but one of the MSBS items, indicating the scale performs similarly for both men and women. Analyses of translations of the MSBS using a Chinese sample {LIU, 2013 #1363} and Spanish sample {Alda, 2015 #995} have helped to confirm the factor structure of these versions of the MSBS. These studies also found a five factor structure fit the MSBS. Alda and colleagues also found good measures of model fit (e.g., CFI = .95) for a model with a second-order 'boredom' factor and the original five first-order factors/subscales. The analyses conducted to date, therefore, suggest the MSBS is a valid scale with good psychometric qualities.

Although the MSBS has shown promising scale characteristics in its early testing, the evidence base for the scale's psychometric utility is still relatively small. In particular, the full English MSBS has yet to receive factor analysis by researchers outside the original authors (i.e. independent factor analysis), and the test-retest reliability of the English MSBS has not been measured. Scores on the MSBS have yet to be compared to scores on another measure of state boredom, the State Boredom Measure (Todman, 2013). Analysis of the Short Form of the MSBS and MSBS-15 has only just begun, so details on the psychometric qualities of

these shortened scales are scarce. The MSBS and its condensed versions appear to be useful measures of state boredom, but additional testing will provide further evidence of the scales' strengths and weaknesses.

Goals of this study

The primary goals of this study were to assess the psychometric utility of the MSBS, its Short Form, and the MSBS-15. A variety of statistical analyses were performed to help achieve these goals. These included a confirmatory factor analysis (CFA) on the full MSBS, assessing the original factor structure identified by Fahlman et al. (2013). Internal consistency, test-retest reliability, and convergent validity of the MSBS, Short Form, and MSBS-15 were also assessed. When assessing the CFA model fit indices, we used guidelines from various authors. These included {Jöreskog, 1993 #991}(recommending smaller chi-square values), {Schumacker, 2004 #997} (recommending GFIs near .95), {Hu, 1999 #977} (recommending CFIs > .95, SRMRs < .08, and RMSEAs < .06), and {Browne, 1993 #990} (recommending RMSEAs < .05).

It should be noted that it was beyond the scope of this study to explore the MSBS from a theoretical perspective. Rather, this study was designed to examine the MSBS with a focus on its statistical properties. The authors acknowledge that a good scale should be supported by both theoretical and statistical evidence. As such, the analyses in this paper should be considered in addition to the strong theoretical support for the MSBS demonstrated by Fahlman et al. (2013).

Method

Participants

The primary sample was a general population adult sample in Australia (original N = 2017, final N = 1716). Participants were recruited via an online paid survey site which rewarded participants for completing the survey. Participants were paid approximately $2-4

for completing the survey. All participants were aged between 18-64. Note that all statistical analyses relate exclusively to this primary sample, unless otherwise stated.

In addition to the primary sample, two smaller samples were collected. An 'advertisement sample' (original $N = 241$, final $N = 199$) was recruited via media releases (e.g., newspaper, radio broadcasts). These participants were entered into a draw for five $100 gift vouchers in return for participating. They were also offered a quick 'boredom summary' at the end of the survey, which provided them with some very basic information about their level of boredom. All participants were aged between 18-64.

A sample of Australian undergraduate psychology students was also assessed (baseline $N = 422$, time two $N = 151$). These students completed the MSBS (and some other surveys) on two occasions with a 6-8 day gap between measurements. Students were awarded course credit as an incentive for their participation. Participants were aged between 16-59.

Demographics

The primary sample for this research consisted of 53.5% females ($n = 918$). Participants in this study ranged from 18 years to 64 years of age, $M = 42.55$, $SD = 13.40$. The participants were primarily Christian (43.2%) or non-religious (42.5%), with 13.8% reporting a different religion. The marital status of the participants varied, with 45.2% being married, 27.7% single or never married, 14.3% living in a de facto relationship, 10.5% being separated or divorced, and 2% having a different marital status. The participants' highest level of education ranged widely, with 31.8% having vocational qualifications, 23.9% having an undergraduate degree, 19.1% having completed year 12, 14.2% completing year 10, 9.8% having postgraduate qualifications, and 1% not completing high school. The annual income of the participants also varied, with 29.5% of respondents earning less than $20 000 per year, 18.5% earning $20 000 - $39 999, 15.2% earning $40 000 - $59 999, 11.2% earning $60 000 - $79 999, 6.5% earning $80 000 - $99 999, 6.8% earning at least $100 000, and 12.4% not

providing income details.

Procedure

All participants completed a range of questionnaires and demographic questions in a single online survey (approx 120 items total). Demographic questions included a range of items based on items used in the Australian national census (Australian Bureau of Statistics, 2011). Some additional items related to participants' experiences of psychiatric disorders, and were not analysed in this study. The survey was completed online at participants' convenience.

Scales

The survey included the Multidimensional State Boredom Scale (MSBS; Fahlman et al., 2013), as well as other scales assessing boredom and mood disturbance. The first two sections of the survey were the Demographic items, followed by the MSBS. The rest of the survey scales were administered in the order described below.

Zuckerman's Boredom Susceptibility Scale. Zuckerman's Boredom Susceptibility Scale (ZBS; Zuckerman, 1994) assesses a person's tendency to become bored (trait boredom). The ZBS has gone through a number of format changes, with Form V being the most frequently employed version. Form V of the ZBS was used in this study. This scale contains 10 forced-choice items, with each item being scored 0 or 1. Higher scores reflect a greater susceptibility to boredom. Although the ZBS has been extensively used in the boredom research literature, its internal consistency is quite poor (e.g., $\alpha = .31$ - .48, Fortune & Goodie, 2010; $\alpha = .44$, Gerritsen, Toplak, Sciaraffa, & Eastwood, 2014).

Depression Anxiety Stress Scales. The Depression Anxiety Stress Scales (DASS; (Lovibond & Lovibond, 1996) are a set of three subscales, measuring depression, anxiety, and stress (as experienced within the last week). The DASS comes in a 21 and 42 item version: The 21 item version was used in this study. Each item is scored from 1-4 (and then

doubled), such that higher scores reflect greater levels of the negative emotional state being assessed. There is good evidence for the utility of the DASS, including high internal consistency for each of the subscales, and a clear factor structure (Antony, Bieling, Cox, Enns, & Swinson, 1998).

Boredom Proneness Scale. The Boredom Proneness Scale (BPS; Farmer & Sundberg, 1986) is a 28 item questionnaire that assessed a person's predisposition to feeling bored (trait boredom). The questionnaire was developed using a true-false format, but many authors (e.g., Vodanovich & Kass, 1990) have adapted the scale to a likert format. In the present study, the BPS was administered using a 7-point likert scale. Higher scores on the BPS reflect a stronger proneness to boredom. Different subscales for the BPS have been derived by multiple authors, with this research including Vodanovich and Kass' five subscales: Affective Responses (five items), Constraint (two items), External Stimulation (eight items), Internal Stimulation (nine items), and Perception of Time (four items). The BPS has considerable evidence of acceptable internal consistency (e.g., Fahlman, Mercer, Gaskovski, Eastwood, & Eastwood, 2009; Wallace, Vodanovich, & Restino, 2003), although its factor structure does not appear to be consistent (Melton & Schulenberg, 2009; Struk, Carriere, Cheyne, & Danckert, 2015).

State Boredom Measure. The State Boredom Measure (SBM; Todman, 2013) is a set of eight test items designed to assess a person's boredom in the preceding week (state boredom). Each test item is measured on a 7-point likert scale, with higher scores reflecting more frequent and/or severe boredom. In contrast to the MSBS, the SBM focuses more on the frequency of boredom, rather than the actual experience and cognitive and emotional correlates of boredom. Additionally, the SBM assesses boredom 'over the last week' (past-state boredom), rather than boredom experienced at the exact time of survey completion (present-state boredom). Unlike many scales, Todman suggests that the SBM items should be

assessed independently - i.e. not summed to create a total scale score. There is evidence of convergent validity for the SBM items (Todman, 2013), with the SBM items being positively related to scores on the BPS and a depression scale.

Screener Question

A 'Screener' question (Berinsky, Margolis, & Sances, 2014) was included within the survey to help identify participants who responded with insufficient effort (Huang, Curran, Keeney, Poposki, & DeShon, 2012). This Screener question was hidden within the BPS items. The question asked the respondents to "Please report 'Neutral' for this item". In total, 205 participants didn't select 'Neutral' for this item, and were removed from further data analyses.

Temporal Discrimination Questions

The administered survey included a relatively large range of questions relating to boredom. As different questions related to different periods of time (i.e., trait boredom vs past-state boredom vs present-state boredom), there was some concern that participants may not properly differentiate between questions that assessed different time periods. For example, participants might have answered all of the scales as if they were referring to 'the past week' rather than the actual time periods described in the instructions for each scale. As such, six items were included in the survey that more explicitly assessed different time periods (e.g., "How frequently have you been bored over the past 30 minutes?", followed by "How frequently have you been bored over the past week?"). These temporal discrimination questions were administered in two sets of three questions, on separate pages of the survey (after the ZBS, and after the BPS), to encourage participants to read the wording of each question. Assessment of these six items was intended to help determine whether participants were responding to the other survey items in an appropriate manner. Specifically, it was expected that correlations between the state boredom temporal discrimination questions

should have higher correlations with the MSBS (state boredom) than the BPS (trait boredom). Conversely, the trait boredom temporal check questions were expected to have higher correlations with the BPS (trait boredom) than the MSBS (state boredom). Such correlations would indicate that respondents were responding to the state (MSBS) and trait (BPS) measures of boredom similarly to the temporal discrimination questions of the corresponding type. Therefore, the relative magnitude of the correlations (in the directions mentioned above) could be used to help detect whether participants were responding to the scales with the correct time period (e.g., present state vs trait).

Results

Data Cleaning and Screening - Primary Sample

A Missing Value Analysis showed that no questions in the MSBS had more than 2% missing data. When data were missing, the (sub)scales of the MSBS were derived by multiplying the existing numbers by the appropriate factor to get the (sub)scale score. For example, if a participant provided responses on only 27 of the 29 MSBS items, the mean of these 27 items was multiplied by 29 to generate the total score for that participant. This score was calculated only if the participant had at least 75% of the data for a (sub)scale; participants with less than 75% data on a (sub)scale did not get a score for that (sub)scale. Given that factor analysis requires no missing data, the Estimisation Maximisation technique was used to impute missing data. This imputed data was only used for the factor analyses reported below - all other analyses use raw data only.

Tests of normality showed no evidence of excessive (> 1 or < -1) skew on any of the MSBS items or its (sub)scale scores. There was, however, mild negative kurtosis (kurtosis values between -1 and -1.2) on 10 of the 29 individual test items. There were no univariate outliers (≥ 3 Interquartile Ranges from the mean) amongst the individual MSBS items or the (sub)scale scores.

In addition to the 205 participants who failed the Screener question, 96 participants were excluded for other reasons. Eighty seven participants who showed evidence of response sets (e.g., responding 1,2,3,4,5,4,3,2,1, or 3,3,3,3,3,3,3,3,3,3) or other potentially dubious data were excluded. Three participants aged under 18 or over 64 (i.e. ineligible for the research) were excluded, as were six participants who reported a poor understanding of English. Analysis of multivariate outliers via inspection of Mahalanobis' Distance values indicated that there were 154 participants with Mahalanobis' D, $p < .001$. It is unclear why so many participants were identified as outliers using this method. A CFA conducted using these outliers removed yielded similar results to a CFA with the outliers retained. Given that there were no obvious problems with these participants' data, their data were retained for the analyses reported below. After all of the data omissions, the original sample of 2017 participants was reduced by 14.92% to a final sample of $N = 1716$, including 790 males and 918 females (eight participants did not report their gender).

Data Cleaning and Screening - Advertisement Sample

Data evaluation procedures similar to those used for the primary sample were performed on the advertisement sample. There was mild negative kurtosis (kurtosis between -1 and -1.4) on 23 of the 29 MSBS items, and the Low Arousal subscale also showed negative kurtosis (kurtosis = -1.009). From the original sample ($N = 241$), 42 were excluded due to concerns over validity. The final sample was $N = 199$, comprising 73 males and 126 females.

Data Cleaning and Screening - Student Sample

Data evaluation procedures similar to those used for the previous samples were performed on the student sample's baseline data and time two data. There was evidence of mild negative kurtosis at baseline, with 18 of the 29 test items having kurtosis values between -1 and -1.35. At time two, there was again mild negative kurtosis observed (10/29 items with kurtosis between -1 and -1.3). Twenty five students' baseline data were excluded due to

validity concerns, as were five students' time two data. The final sample size for the baseline data was 422, which included 335 females and 87 males. The final sample size for the time two data was 151, which comprised 116 females and 35 males.

Descriptive Statistics

Descriptive statistics for each of the scales used in the research are provided in Table 1.

Temporal Discrimination Questions

Six test items were used to help determine whether participants responded differentially to the different time periods intended to be assessed by the various scales in the study. See Table 2 for correlational data for these items with the MSBS Full Scale, MSBS Short Form, MSBS-15, and BPS. The data indicate that participants' responses to these six test items were relatively similarly related to their state (MSBS) and trait (BPS) scale scores.

Assessing the Suitability of Factor Analysis

Analysis of the correlation matrix indicated that each of the MSBS variables was positively related to the other variables, with most correlations between 0.3 - 0.7. A copy of this matrix is available from the first author. There were three instances in which an inter-correlation was above .8, but none of these were above .9. Bartlett's Test of Sphericity for the primary sample was significant at < .001. Similarly, the KMO measure of sampling adequacy was 0.978, suggesting that the primary sample had excellent sampling adequacy.

The determinant of the correlation matrix was below .0001, suggesting possible multicollinearity (Tabachnick & Fidell, 2001). All Variance Inflation Factors (VIF) were below 10. However, two variables had a VIF higher than 5 (Item 18: 5.714, and Item 26: 5.239) These items read: (item 18) "Time is moving very slowly"; and (item 26) "Right now it seems like time is passing slowly". Four items had a variance proportion above 0.5: Item 11 (.55), item 18 (.78), item 24 (0.52), item 26 (.68). There was a single condition index greater than 30 observed, and most of the condition indices were above 15. Given these metrics, it

was determined that potentially disruptive multicollinearity was present in the data, likely within items 18 and 26.

Confirmatory Factor Analyses

A series of confirmatory factor analyses (CFAs) were run on the data using AMOS 22.0.0 (Arbuckle, 2013). These models were run using maximum likelihood estimation. Table 3 contains model fit indices for each of the CFAs run on the full MSBS. Table 4 contains the fit indices for the MSBS-15, while Table 5 contains fit indices for the MSBS Short Form. More detailed copies of these CFAs (including factor loadings) are available from the first author.

MSBS Full Scale. The first specified model was based on Fahlman et al.'s (2013) structural model of the MSBS. This model includes one second order factor - Boredom. Connected to Boredom, there are five first order factors: Disengagement, High Arousal, Inattention, Low Arousal, and Time Perception. These five factors contain 10, five, four, five, and five test items, respectively. The test items connected to each of the factors were the same as those identified in Fahlman et al. See Figure 1 for a pictorial representation of the specified model.

After running the initial CFA and examining the goodness of fit indices obtained, it was decided that the specified model had a decent to good fit to the data. The goodness of fit indices obtained in this CFA were somewhat worse than those obtained by Fahlman et al., (2013). An alternate model was tested that combined the high and low arousal items into a single factor, as per Fahlman and colleagues' (2013) testing. This analysis yielded similar, but slightly worse, goodness of fit indices than the original model. A model containing just the five dimensions of boredom - but no second order factor - was similarly fit to the data as the original model. A model with only one factor was poorly fit.

The original model from Fahlman et al. (2013; see Figure 1) was also tested on the

advertisement sample ($N = 199$) and the student baseline data ($N = 422$). For the advertisement sample, somewhat worse fit was achieved relative to the primary sample. For the student sample, the analysis showed a decent fit for the model.

MSBS-15. A similar series of analyses to those run on the full MSBS were conducted for the MSBS-15. The first CFA assessed a model similar to that of Fahlman et al. (2013), although containing only the 15 items of the MSBS-15. This model had good fit to the data. Combining the arousal items resulted in a slightly worse fit. Removing the second order 'boredom' factor showed good fit, but a single factor model had poor fit. Testing the 'original' model (based on Fahlman and colleagues' model but with only 15 items) for the advertisement and student samples resulted in decent fit and good fit, respectively.

MSBS Short Form. Analyses on the MSBS Short Form only assessed a single factor model. This scale does not contain items from all of the original MSBS subscales, and a reading of Hunter et al. (2015) implies that the Short Form is not intended to capture five dimensions of boredom. The single factor CFA on the primary sample yielded decent fit. The advertisement sample showed a mediocre fit for the model, while the student sample had mediocre-decent fit.

Internal Consistency

Internal consistency of the MSBS and its subscales was assessed for the total sample, as well as for males and females separately. Similarly, the MSBS Short Form and MSBS-15 were assessed. Table 6 shows that each of the subscales of the MSBS, the full MSBS, the Short Form, and the MSBS-15, all demonstrated very high internal consistency, in both the male and female samples. There was little difference in the consistency between the genders.

Convergent Validity

Convergent validity of the full MSBS and each of its subscales, Short Form, and MSBS-15, was assessed via comparisons with the other measures used in the study (see Table 7).

The intercorrelations of the MSBS (sub)scales were also calculated. The MSBS showed only weak-moderate positive correlations with the ZBS. The MSBS demonstrated moderate to very strong positive correlations with the BPS and each of its subscales. The MSBS was moderately-strongly positively related to all of the SBM items. The MSBS was strongly and positively related to depression, anxiety, and stress, as measured by the DASS. The results for the MSBS Short Form and MSBS-15 were very similar to those for the full scale, with the correlations observed generally only slightly weaker for the Short Form and MSBS-15. The correlation between the Short Form and the full MSBS was extremely high (r = .962), and the MSBS-15 was also very highly correlated with the full MSBS (r = .988). The MSBS (sub)scale intercorrelations were all strong or very strong.

Test-retest Reliability

Test-retest reliability of the full MSBS, Short Form, and MSBS-15 were assessed using data from the student sample (time two N = 151). Over the 6-8 days between measurements, students showed high consistency in their scores. The test-retest correlations for the MSBS (sub)scales were: Disengagement (r = .659), High Arousal (r = .634), Inattention (r = .663), Low Arousal (r = .695), Time Perception (r = .597), Full Scale MSBS (r = .708), Short Form (r = .694), MSBS-15 (r = .680). For comparison, test-retest reliability of the BPS was r = .811. All correlations were significant at $p < .001$.

Discussion

The preceding data analyses have suggested several strengths and weaknesses of the Multidimensional State Boredom Scale. The convergent validity, internal consistency, and test-retest reliability of the MSBS and Short Form were all very good. Factor analyses confirmed that Fahlman's (2013) model was the best fitting model for the MSBS. However, there is some evidence that the MSBS subscales may not consistently measure particularly distinct concepts.

Factor Structure of the MSBS

The factor analyses and other data analyses on the MSBS led to some interesting - and somewhat ambiguous - results. The CFAs using Fahlman and colleagues' (2013) original model suggested a decent to good fit to the data in the primary sample (e.g., GFI = .855, CFI = .929). Attempts to improve the model, including combining the arousal items, removing the second order factor, and using a single factor structure, did not substantially improve the fit of the model. Taken as a whole, the CFAs indicate that the originally proposed model for the MSBS is the best fitting model for the scale.

There was an almost perfect relationship between the full MSBS and the Short Form (r = .962) that consists largely of 'Disengagement' test items. Similarly, the full MSBS had a very strong relationship with the 'Disengagement' subscale itself (r = .962). This indicates that the full MSBS assesses the disengagement component of boredom very well, but there is some ambiguity as to the distinctness/uniqueness of the other MSBS dimensions.

The intercorrelations of the MSBS subscales were generally very high (r = .574 to .860). The Time Perception subscale had weaker intercorrelations, but even these intercorrelations were all above .5. Similarly, most of the inter-item correlations were moderate-strong. The correlations between the MSBS subscales and other boredom measures (e.g., the BPS) were similar in magnitude regardless of the MSBS subscale being assessed, except for the Time Perception subscale correlations. All of these findings suggest that the MSBS may not consistently assess especially separated components of boredom. That is, the MSBS subscales of disengagement, high arousal, inattention, low arousal, and (to a lesser extent) time perception, may be assessing relatively similar constructs. It should be kept in mind, however, that strong subscale intercorrelations may reflect very strong relationships between the latent constructs being assessed by the MSBS. That is, disengagement, high arousal, inattention, low arousal, and time perception problems may be inherently strongly related,

which means any self-report measures of these constructs will ideally produce strong intercorrelations.

Baratta and Spence (2015) raise an important issue about the use of full-scale scores for multidimensional scales. Baratta and Spence argue that using a total score for a multidimensional boredom scale makes the assumption that all test items are parallel {Graham, 2006 #1358}. For test items to be parallel, they should (among other things) measure the same construct (i.e. be unidimensional), and have the same 'difficulty' (i.e. indicate the same degree of the measured construct) {DeVellis, 2006 #1362}. Unfortunately, neither of these assumptions may be true for the MSBS. Previous theory and statistical analyses of the MSBS have indicated that the MSBS captures multiple dimensions of boredom (e.g. {Fahlman, 2013 #476} {Baratta, 2015 #1004}. {Baratta, 2015 #1004} used a polytomous item response theory analysis that demonstrated that the MSBS items have different difficulties. As such, one can argue that the full scale score for the MSBS may not always be completely reliable and/or valid as a single metric for state boredom.

Factor Structure of the MSBS-15

The present study's MSBS-15 CFA results are similar to the results obtained by Baratta and Spence (2015). Specifically, the factor analyses of the MSBS-15 in this paper suggest a multidimensional structure for the scale (e.g., primary sample GFI = .919, CFI = .961). Similarly, when Baratta and Spence compared several structural models on the MSBS-15, their results indicated a clear superiority of their multivariate and superordinate models over a unidimensional model for the MSBS-15. The results of the present study support the use of the MSBS-15, and actually suggest that it may capture the dimensions of boredom more distinctly than the full scale. This scale is currently recommended for capturing multiple dimensions of state boredom with a relatively minimal participant burden.

Factor Structure of the MSBS Short Form

The CFAs on the MSBS Short Form indicated mediocre to decent fit for a single factor model (e.g., primary sample GFI = .933 and CFI = .911, but RMSEA = .114). However, given that the Short Form contains five items from the original Disengagement subscale, a single item from the Time Perception subscale, only two items from the Inattention subscale, and no items relating to low or high arousal, it is not clear what an alternate model for this scale would consist of. Exploratory factor analyses were performed, and strongly indicated the presence of just one factor (for example, the second Eigenvalue was 0.780; see Appendix for details). As such, the Short Form appears to be best thought of as a unidimensional measure of boredom, albeit one that includes some items that do not load as strongly on this factor as would be preferred. It should be noted that the Short Form was developed with the intention of accurately classifying people as either bored or not bored {Hunter, 2015 #975}, rather than being intended to fully capture a unidimensional construct. As such, the model fit indices obtained in this research do not suggest that the scale is not suitable for its intended purposes. Rather, one must simply consider the scale's strengths and weaknesses when considering it for inclusion in a study (see{Hunter, 2015 #975} for related discussion).

Internal Consistency

The internal consistency figures for the MSBS (sub)scales were all very high. This finding is relatively consistent with the results of previous analyses of the MSBS (Fahlman et al., 2009; Fahlman et al., 2013; Goldberg, Eastwood, Laguardia, & Danckert, 2011). Importantly, the data also demonstrate similarly high consistency figures for men and women.

Convergent Validity

The MSBS, its subscales, and the MSBS condensed scales were all strongly related to the BPS and the Depression Anxiety Stress Scales. The correlation observed between the MSBS total score and the Depression subscale of the DASS (r = .782) was perhaps higher than one

would want from two theoretically distinct concepts. However, the correlation observed was not high enough to indicate that the MSBS actually measures depression rather than boredom, and previous research has indicated a separation between the MSBS and depression (Goldberg et al., 2011). The correlations between the MSBS and the SBM items were moderate to strong, which provides good evidence of construct validity for the MSBS. The ZBS was only moderately related to the MSBS (sub)scales, but previous studies have observed similar correlations between the ZBS and other boredom measures (e.g., Farmer & Sundberg, 1986; Gerritsen, Goldberg, & Eastwood, 2015; Mercer-Lynn, Bar, & Eastwood, 2014). Overall, the MSBS and its condensed versions demonstrated good convergent validity in this study, with strong correlations observed for theoretically related scales/constructs (boredom proneness, depression, anxiety, and stress). Future research could continue to explore the convergent validity of the MSBS and its condensed versions with other relevant constructs, such as aggression (Dahlen, Martin, Ragan, & Kuhlman, 2004), disinhibition (Fortune & Goodie, 2010), and hyperactivity (Gerritsen, Toplak, Sciaraffa, & Eastwood, 2014).

Test-retest Reliability

The test-retest reliability of the MSBS over a period of about one week was moderately high (r = .708 for the full scale, r = .694 for the Short Form, r = .680 for the MSBS-15). The test-retest reliability of the BPS (a trait measure of boredom) was .811. These figures are fairly similar to those observed for a measure of anxiety: The Spielberger State-Trait Anxiety Inventory. A review by Barnes, Harp, and Jung (2002) on this anxiety scale indicated that the state version had a mean test-retest reliability of .70, compared with .88 for the trait version. Overall, the data suggest that the MSBS exhibits fairly high consistency over a one week period.

The test-retest reliability correlations obtained in this study are smaller than those

obtained using a Spanish version of the MSBS {Alda, 2015 #995}. Alda and colleagues observed test-retest reliabilities over a 1-2 week period, with the total Spanish MSBS score having an autocorrelation of .90. If the test-retest reliability of the English MSBS over 1-2 weeks was as high as .90, this would raise concerns about whether participants are responding too consistently on the scale, given the expected daily fluctuations in people's state boredom.

It is hoped that future research examines the test-retest reliability of the MSBS over both much longer periods of time, as well as very short periods of time. Very short test-retest periods (e.g., within 1 hour) would allow for the measurement of the scale's dependability. Dependability refers to a scale's autocorrelation when the timing between tests is not long enough for the participant to actually change on the assessed variable (Cattell, Eber, & Tatsuoka, 1970). This type of measurement allows for researchers to verify whether the scale exhibits high consistency (and low measurement error).

Analyses of MSBS scores over a period of months or years would help to establish the long term stability of the scale. According to Cattell et al. (1970), stability refers to the autocorrelation of a scale when the timing between tests is two or more months. Given the contrast with dependability, it seems reasonable to instead consider stability as a scale's autocorrelation when the timing between tests is long enough that the participant has probably changed on the concept being measured (see Watson, 2004, for related discussion). That is, if a construct such as state boredom is likely to change over a given time period (e.g., one week), the scale's autocorrelation could be considered as a measure of stability. The present study's test-retest reliability for the MSBS (.708) could therefore be considered as a (relatively short-term) measure of stability.

It seems likely that state boredom stability may be different for people with different levels of trait boredom. That is, a person with high boredom proneness may have greater

consistency in their levels of state boredom. The highly boredom prone individual may frequently feel state boredom regardless of the situations they are in (see Fisher, 1993; Hamilton, 1981; Mercer-Lynn, Bar, & Eastwood, 2014). In contrast, someone with low boredom proneness may be more influenced by situational factors, resulting in inconsistent levels of state boredom. Longitudinal studies that assess the MSBS, trait boredom, and situational factors, could provide additional information about the relative influences of a person's situation and their trait boredom on their state boredom.

Temporal Discrimination - Differentiating State and Trait Boredom

The temporal discrimination questions assessed in this research cast some doubt on whether participants responded appropriately to the different time periods being assessed by different scales in this research. If participants correctly discerned the time periods being assessed by the temporal discrimination items, MSBS items, and BPS items, one might expect there to be substantially different correlations observed when comparing the temporal discrimination items to state and trait measures of boredom. That is, temporal discrimination items that explicitly assess state boredom (e.g., "How intense is your boredom right now?") should be more strongly related to the MSBS (which measures state boredom) than to the BPS (which measures trait boredom). Conversely, temporal discrimination items that assess trait boredom (e.g., "How intense has your boredom usually been during the last five years?") should be more strongly related to the BPS than to the MSBS.

The data obtained showed that the trait boredom temporal discrimination items were generally only slightly more strongly related to the BPS than to the MSBS. This was especially true of the primary sample. Additionally, the correlations observed between the state boredom temporal discrimination items and the MSBS were generally only slightly larger than those for the BPS. The data suggest that participants may not have correctly distinguished between trait and state boredom when answering some test items (either the

temporal discrimination items, and/or the MSBS and/or BPS items).

The very strong correlation between the MSBS Full Scale and BPS (r = .815) in this study is unusually high. Previous research has shown slightly weaker correlations between these two measures: r = .62 (Fahlman et al., 2013), r = .711 (Goldberg et al., 2011), r = .712 (Gerritsen et al., 2014). The high correlation observed in this study may partly reflect insufficient effort responding (Huang et al., 2012) in this sample, or it may reflect a genuinely strong congruency between the BPS and MSBS constructs.

It is hoped that future research can help to determine whether the participants in this research exhibited relatively poor temporal discrimination when answering questions, and/or how strongly state and trait boredom are related. It is recommended that authors of future studies make particular efforts to ensure that participants understand that different questionnaires relate to different time periods. The use of non-survey methods of assessing boredom may prove fruitful in helping to assess the relationship between state and trait boredom (see Vogel-Walcutt, Fiorella, Carper, & Schatz, 2012, for discussion of non-survey assessments of boredom).

Strengths and Limitations of this Research, and Areas for Future Study

The research allowed for a wide range of statistical analyses to be performed on the MSBS Full Scale, MSBS-15, and Short Form. These analyses have helped to elucidate the nature of the MSBS and its relationship to other measures of mood. One strength of the research was the use of three samples to aid in the CFA process. The use of three separately recruited samples allowed for multiple verifications of Fahlman and colleagues' (2013) structural model. The primary sample and student sample were both relatively large (N = 1716 and N = 422, respectively), which adds further support to the CFA procedures employed.

An additional strength of the study was the range of procedures used to screen

participants' data validity. The Screener question helped to identify about 10% of the primary sample as having potentially invalid data. This is a substantial portion of the data, and provides further support for the use of such screening procedures. Unfortunately, the temporal discrimination questions data suggest that the remaining participants may not have adequately discriminated between the different time periods measured by the questionnaires used in this research. It is hoped that future research employs more detailed investigations of participants' temporal discrimination when concurrently assessing state and trait boredom.

A significant limitation of the research design was the likelihood of the context of assessment influencing scores on the MSBS. Specifically, participants completed the survey online at their leisure. As such, their level of state boredom at the time of measurement would largely be influenced by the activity of completing the survey itself. Given that the scale was relatively long (about 120 items) and featured many items assessing similar constructs, this may have induced boredom among the participants. The MSBS was completed early in the survey, but nonetheless, MSBS scores in this study may largely reflect boredom caused by doing surveys. If completing surveys is an especially boring task (which for many people, it would be), then this would lead to MSBS scores in this study reflecting 'survey-induced boredom'. As such, when the MSBS is completed in other contexts (e.g., as a clinical instrument), the data may not be directly comparable to the data in this study. Accordingly, the data analyses and conclusions of this paper should be considered as relevant to survey research contexts rather than clinical uses of the MSBS. For example, it is entirely possible that the MSBS would have very different test-retest reliabilities and/or a different factor structure when used as a clinical assessment tool.

A potential limitation of the research was the use of a paid sample for the primary sample. It is possible that the results for this sample may not be generalisable to the general population. There is some evidence from the data that this sample were especially prone to

insufficient effort responding (Huang et al., 2012), including the extent of dubious data observed, and the temporal discrimination difficulties discussed above. Given budget constraints, it is hoped that future research recruits participants from a range of sources to help reduce potential sample biases.

Conclusion

This research has helped to verify or establish several important psychometric characteristics of the Multidimensional State Boredom Scale and its condensed versions. The convergent validity of the MSBS, Short Form, and MSBS-15 appears to be quite good. Similarly, the internal consistency of these scales is excellent. Test-retest reliability of the scales was also high. Factor analyses supported the factor structure for the MSBS originally proposed by Fahlman et al. (2013). However, the analyses conducted suggest that the boredom dimensions assessed by the MSBS may not be well separated concepts. It is hoped that future research examines this issue in more detail. This study provides further evidence that the MSBS is the state boredom measure with the most evidence for its construct validity, convergent validity, test-retest reliability, and internal consistency. As such, the MSBS (or preferably, the MSBS-15) are recommended for boredom researchers that wish to assess state boredom.

References

Ahmed, S. M. S. (1990). Psychometric properties of the Boredom Proneness Scale. *Perceptual and Motor Skills, 71*(3), 963-966. doi: 10.2466/pms.1990.71.3.963

Alda, M., Minguez, J., Montero-Marin, J., Gili, M., Puebla-Guedea, M., Herrera-Mercadal, P., . . . & Garcia-Campayo, J. (2015). Validation of the Spanish version of the Multidimensional State Boredom Scale (MSBS). *Health and Quality of Life Outcomes, 13*(1). Advance online publication. doi: 10.1186/s12955-015-0252-2

Antony, M. M., Bieling, P. J., Cox, B. J., Enns, M. W., & Swinson, R. P. (1998). Psychometric properties of the 42-item and 21-item versions of the Depression Anxiety Stress Scales in clinical groups and a community sample. *Psychological Assessment, 10*(2), 176-181. doi: 10.1037/1040-3590.10.2.176

Arbuckle, J. L. (2013). AMOS (version 22.0.0; Computer software). Meadville, PA: IBM Corporation.

Australian Bureau of Statistics. (2011). *National Census Household Form.* Retrieved from http://www.abs.gov.au/websitedbs/censushome.nsf/home/2011hhftranscript%5C

Baratta, P. L., & Spence, J. R. (2015). A riddle, wrapped in a mystery, inside an enigma . . . or just multidimensional? Testing the multidimensional structure of boredom. In N. M. Ashkanasy, W. J., Zerbe, & C. E. J. Härtel (Series Eds.), *Research on Emotion in Organisations: Vol. 11. New ways of studying emotions in organizations* (pp. 139-172). Emerald Group.

Barnes, L. L. B., Harp, D., & Jung, W. S. (2002). Reliability generalization of scores on the Spielberger State-Trait Anxiety Inventory. *Educational and Psychological Measurement, 62*(4), 603-618. doi: 10.1177/0013164402062004005

Bensley, L., Nelson, N., Kaufman, J., Silverstein, B., & Shields, J. W. (1995). Patient and staff views of factors influencing assaults on psychiatric hospital employees. *Issues in*

Mental Health Nursing, 16(5), 433-446. doi: doi:10.3109/01612849509006944

Berinsky, A. J., Margolis, M. F., & Sances, M. W. (2014). Separating the shirkers from the workers? Making sure respondents pay attention on self-administered surveys. *American Journal of Political Science, 58*(3), 739-753. doi: 10.1111/ajps.12081

Blaszczynski, A., McConaghy, N., & Frankova, A. (1990). Boredom proneness in pathological gambling. *Psychological Reports, 67*(1), 35-42. doi: 10.2466/PR0.67.5.35-42

Browne, M., & Cudeck, R. (1993). Alternative ways of assessing model fit. In K. A. Bollen & J. S. Long (Eds.), *Testing structural equation models* (pp. 136-162). Newbury Park, CA: Sage.

Bruursema, K., Kessler, S. R., & Spector, P. E. (2011). Bored employees misbehaving: The relationship between boredom and counterproductive work behaviour. *Work and Stress, 25*(2), 93-107. doi: 10.1080/02678373.2011.596670

Cattell, R. B., Eber, H. W., & Tatsuoka, M. M. (1970). Handbook for the sixteen personality factor questionnaire (16PF) (pp. 1-388). Champaign, IL: Institute for Personality and Ability Testing.

Chaplin, R., McGeorge, M., & Lelliott, P. (2006). Violence on inpatient units for people with learning disability and mental illness: The experiences of service users and staff. *The British Journal of Development Disabilities, 52*(103), 109-116. doi: 10.1179/096979506799103532

Cortina, J. M. (1993). What is coefficient alpha? An examination of theory and applications. *Journal of Applied Psychology, 78*(1), 98. doi: 10.1037/0021-9010.78.1.98

Dahlen, E. R., Martin, R. C., Ragan, K., & Kuhlman, M. M. (2004). Boredom proneness in anger and aggression: Effects of impulsiveness and sensation seeking. *Personality and Individual Differences, 37*(8), 1615-1627. doi: 10.1016/j.paid.2004.02.016

Fahlman, S., Mercer-Lynn, K., Flora, D. B., & Eastwood, J. D. (2013). Development and validation of the Multidimensional State Boredom Scale. *Assessment, 20*(1), 68-85. doi: 10.1177/1073191111421303

Fahlman, S. A., Mercer, K. B., Gaskovski, P., Eastwood, A. E., & Eastwood, J. D. (2009). Does a lack of life meaning cause boredom? Results from psychometric, longitudinal, and experimental analyses. *Journal of Social and Clinical Psychology, 28*(3), 307-340. doi: 10.1016/j.concog.2005.11.009

Farmer, R., & Sundberg, N. D. (1986). Boredom proneness - the development and correlates of a new scale. *Journal of Personality Assessment, 50*(1), 4-17. doi: 10.1207/s15327752jpa5001_2

Fisher, C. D. (1993). Boredom at work: A neglected concept. *Human Relations, 46*(3), 395-417. Retrieved from http://hum.sagepub.com/

Fortune, E. E., & Goodie, A. S. (2010). The relationship between pathological gambling and sensation seeking: The role of subscale scores. *Journal of Gambling Studies, 26*(3), 331-346. doi: 10.1007/s10899-009-9162-8

Gerritsen, C. J., Goldberg, J. O., & Eastwood, J. D. (2015). Boredom proneness predicts quality of life in outpatients diagnosed with schizophrenia-spectrum disorders. *International Journal of Social Psychiatry*. Advance online publication. doi: 10.1177/0020764015584647

Gerritsen, C. J., Toplak, M. E., Sciaraffa, J., & Eastwood, J. (2014). I can't get no satisfaction: Potential causes of boredom. *Consciousness and Cognition, 27*, 27-41. doi: 10.1016/j.concog.2013.10.001

Goldberg, Y. K., Eastwood, J. D., Laguardia, J., & Danckert, J. (2011). Boredom: An emotional experience distinct from apathy, anhedonia, or depression. *Journal of Social and Clinical Psychology, 30*(6), 647-666. doi: 10.1521/jscp.2011.30.6.647

Hamilton, J. A. (1981). Attention, personality, and the self-regulation of mood: Absorbing interest and boredom. *Progress in Experimental Personality Research, 10*(28), 281-308.

Hamilton, J. A., Haier, R. J., & Buchsbaum, M. S. (1984). Intrinsic enjoyment and boredom coping scales: Validation with personality, evoked potential and attention measures. *Personality and Individual Differences, 5*(2), 183-193. doi: 10.1016/0191-8869(84)90050-3

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

Huang, J. L., Curran, P. G., Keeney, J., Poposki, E. M., & DeShon, R. P. (2012). Detecting and deterring insufficient effort responding to surveys. *Journal of Business and Psychology, 27*(1), 99-114. doi: 10.1007/s10869-011-9231-8

Hunter, J. A., Dyer, K. J., Cribbie, R. A., & Eastwood, J. D. (2015). Exploring the utility of the Multidimensional State Boredom Scale. *European Journal of Psychological Assessment*. Advance online publication. doi: 10.1027/1015-5759/a000251

Inman, A., Kirsh, K., & Passik, S. (2003). A pilot study to examine the relationship between boredom and spirituality in cancer patients. *Palliative and Supportive Care, 1*(2), 143-151. doi: 10.1017/S1478951503030165

Loubser, I., Chaplin, R., & Quirk, A. (2009). Violence, alcohol and drugs: The views of nurses and patients on psychiatric intensive care units, acute adult wards and forensic wards. *Journal of Psychiatric Intensive Care, 5*(1), 33-39. doi: 10.1017/S1742646408001386

Lovibond, S.H. & Lovibond, P.F. (1995). *Manual for the Depression Anxiety Stress Scales* (2nd ed.). Sydney, Australia: Psychology Foundation.

Meehan, T., McIntosh, W., & Bergen, H. (2006). Aggressive behaviour in the high-secure

forensic setting: The perceptions of patients. *Journal of Psychiatric and Mental Health Nursing, 13*(1), 19-25. doi: 10.1111/j.1365-2850.2006.00906.x

Melton, A. M. A., & Schulenberg, S. E. (2009). A confirmatory factor analysis of the Boredom Proneness Scale. *The Journal of Psychology, 143*(5), 493-508. doi: 10.3200/JRL.143.5.493-508

Mercer-Lynn, K. B., Bar, R. J., & Eastwood, J. D. (2014). Causes of boredom: The person, the situation, or both? *Personality and Individual Differences, 56*, 122-126. doi: 10.1016/j.paid.2013.08.034

Newell, S. E., Harries, P., & Ayers, S. (2011). Boredom proneness in a psychiatric inpatient population. *International Journal of Social Psychiatry, 58*(5), 488-495 doi: 10.1177/0020764011408655

Ng, A. H., Liu, Y., Chen, J.-z., & Eastwood, J. D. (2015). Culture and state boredom: A comparison between European Canadians and Chinese. *Personality and Individual Differences, 75*, 13-18. doi: 10.1016/j.paid.2014.10.052

Panayides, P. (2013). Coefficient alpha: Interpret with caution. *Europe's Journal of Psychology, 9*(4), 687–696. doi: 10.5964/ejop.v9i4.653

Schermelleh-Engel, K., Moosbrugger, H., & Müller, H. (2003). Evaluating the fit of structural equation models: Tests of significance and descriptive goodness-of-fit measures. *Methods of Psychological Research Online, 8*(2), 23-74. Retrieved from http://www.dgps.de/fachgruppen/methoden/mpr-online/

Struk, A. A., Carriere, J. S. A., Cheyne, J. A., & Danckert, J. (2015). A Short Boredom Proneness Scale: Development and psychometric properties. *Assessment*. Advance online publication. doi: 10.1177/1073191115609996

Tabachnick, B. G., & Fidell, L. S. (2001). *Using multivariate statistics* (4th ed.). Boston, MA: Allyn and Bacon.

Todman, M. (2013). The dimensions of state boredom: Frequency, duration, unpleasantness, consequences and causal attributions. *Academic Research International, 1*(1), 32-40. Retrieved from http://www.journals.savap.org.pk/

Ullman, J. (2001). Structural equation modeling. In B. G. Tabachnick & L. S. Fidell (Eds.), *Using multivariate statistics* (pp. 676-780). Boston, MA: Allyn & Bacon.

Vodanovich, S. J., & Kass, S. J. (1990). A factor analytic study of the Boredom Proneness Scale. *Journal of Personality Assessment, 55*(1-2), 115-123. doi: 10.1207/s15327752jpa5501&2_11

Vodanovich, S. J., & Watt, J. D. (2015). Self-report measures of boredom: An updated review of the literature. *The Journal of Psychology*. Advance online publication. doi: 10.1080/00223980.2015.1074531

Vodanovich, S. J., Verner, K. M., & Gilbride, T. V. (1991). Boredom proneness: Its relationship to positive and negative affect. *Psychological Reports, 69*(3f), 1139-1146. doi: 10.2466/PR0.69.8.1139-1146

Vogel-Walcutt, J., Fiorella, L., Carper, T., & Schatz, S. (2012). The definition, assessment, and mitigation of state boredom within educational settings: A comprehensive review. *Educational Psychology Review, 24*(1), 89-111. doi: 10.1007/s10648-011-9182-7

Wallace, J. C., Vodanovich, S. J., & Restino, B. M. (2003). Predicting cognitive failures from boredom proneness and daytime sleepiness scores: An investigation within military and undergraduate samples. *Personality and Individual Differences, 34*(4), 635-644. doi: 10.1016/S0191-8869(02)00050-8

Watson, D. (2004). Stability versus change, dependability versus error: Issues in the assessment of personality over time. *Journal of Research in Personality, 38*(4), 319-350. doi: 10.1016/j.jrp.2004.03.001

Wiesbeck, G. A., Wodarz, N., Mauerer, C., Thome, J., Jakob, F., & Boening, J. (1996).

Sensation seeking, alcoholism and dopamine activity. *European Psychiatry, 11*(2), 87-92. doi: 10.1016/0924-9338(96)84785-7

Wiesner, M., Windle, M., & Freeman, A. (2005). Work stress, substance use, and depression among young adult workers: An examination of main and moderator effect models. *Journal of Occupational Health Psychology, 10*(2), 83-96. doi: 10.1037/1076-8998.10.2.83

Woodall, J. (2012). Social and environmental factors influencing in-prison drug use. *Health Education, 112*(1), 31-46. doi: 10.1108/09654281211190245

Zuckerman, M. (1994). *Behavioral expressions and biosocial bases of sensation seeking* (2nd ed.). New York, NY: Cambridge University.

Figure 1

Measurement model used to assess the factor structure of the MSBS, based on Fahlman et al. (2013)

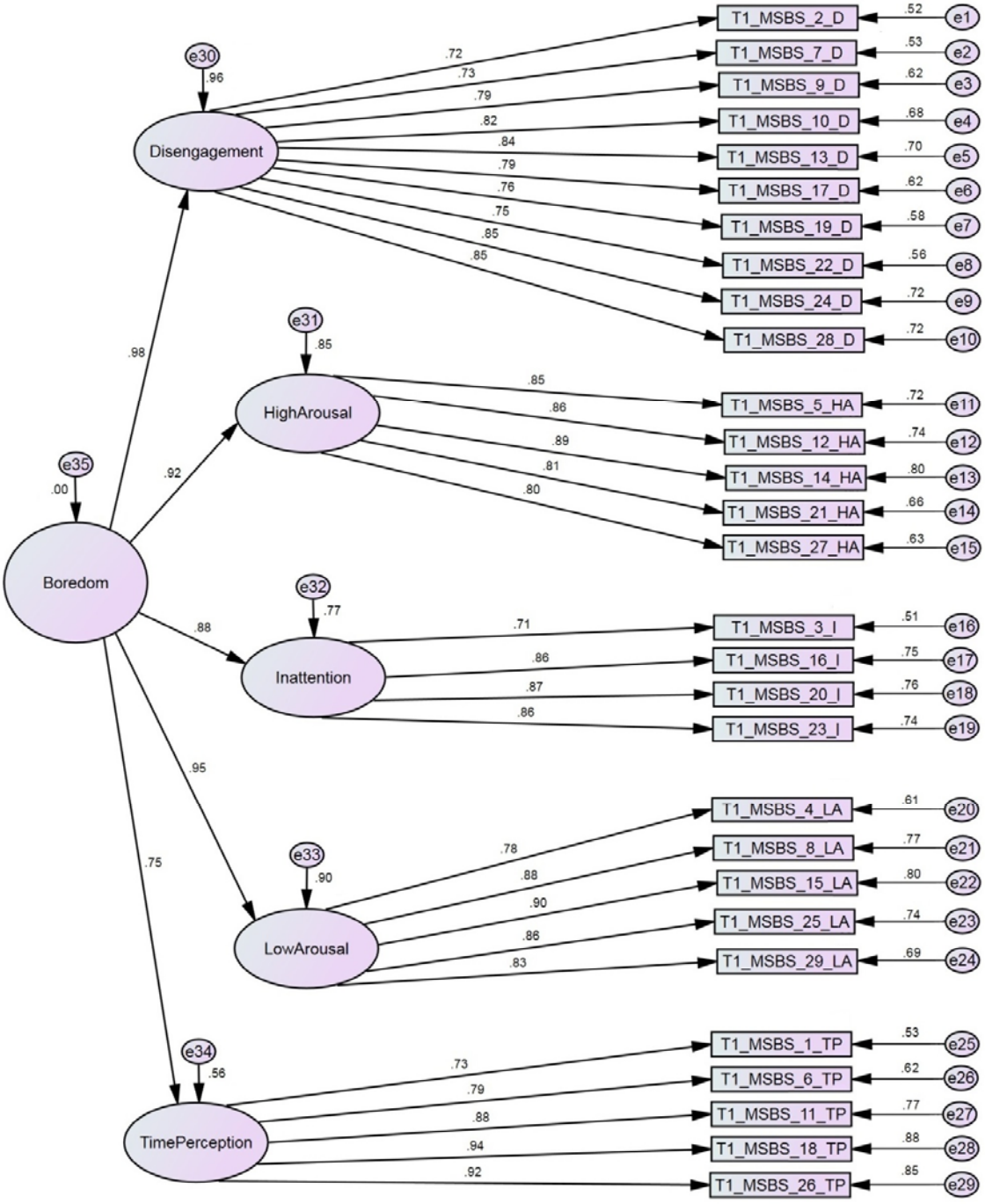

Table 1.

Descriptive statistics for each scale for the primary sample (N ranges from 1691 - 1715)

(Sub)scale	*M*	*SD*	99% CIs
MSBS Short Form	27.84	10.84	27.16 - 28.51
MSBS-15	48.95	21.26	47.62 - 50.27
MSBS Full Scale	96.97	38.75	94.55 - 99.38
MSBS - Disengagement	35.70	14.36	34.81 - 36.60
MSBS - High Arousal	16.41	7.51	15.94 - 16.88
MSBS - Inattention	14.22	5.97	13.85 - 14.59
MSBS - Low Arousal	16.56	8.21	16.04 - 17.07
MSBS - Time Perception	14.07	7.05	13.62 - 14.51
ZBS	2.24	1.81	2.13 - 2.35
BPS Total Score	99.52	22.86	98.10 - 100.94
SBM - Item 1	2.88	1.61	2.78 - 2.98
SBM - Item 2	2.24	1.55	2.14 - 2.34
SBM - Item 3	2.76	1.64	2.66 - 2.86
SBM - Item 4	3.53	1.80	3.42 - 3.64
SBM - Item 5	2.90	1.77	2.78 - 3.01
SBM - Item 6	3.38	1.81	3.26 - 3.49
SBM - Item 7	2.72	1.99	2.60 - 2.85
SBM - Item 8	3.00	1.91	2.88 - 3.12
DASS - Depression	12.07	5.57	11.73 - 12.42
DASS - Anxiety	10.38	4.28	10.11 - 10.65
DASS - Stress	12.24	4.81	11.94 - 12.54

Notes: MSBS = Multidimensional State Boredom Scale; ZBS = Zuckerman's Boredom Susceptibility Scale; BPS = Boredom Proneness Scale, SBM = State Boredom Measure,

DASS = Depression Anxiety Stress Scales, DASS-D = Depression subscale, DASS-A = Anxiety subscale, DASS-S = Stress subscale

Table 2

Correlations between temporal discrimination questions and state and trait boredom

Test item	Primary Sample ($N \approx 1700$)				Advertisement Sample ($N \approx 197$)			
	MSBS Short Form	MSBS-15	MSBS Full Scale	BPS Full Scale	MSBS Short Form	MSBS-15	MSBS Full Scale	BPS Full Scale
How intense is your boredom right now? (state)	.648	.635	.651	.592	.699	.640	.685	.636
How intense has your boredom usually been during the last week? (state)[a]	.684	.685	.695	.644	.677	.653	.696	.672
How intense has your boredom usually been during the last five years? (trait)	.578	.588	.593	.601	.477	.474	.491	.578
How frequently have you been bored over the past 30 minutes? (state)	.571	.549	.568	.552	.602	.526	.573	.578
How frequently have you been bored over the past week? (state)[a]	.685	.674	.691	.680	.696	.669	.709	.695
How frequently have you been bored over the past five years? (trait)	.577	.588	.595	.621	.451	.456	.469	.608

Notes. [a] These items were deemed to be state boredom measures rather than trait boredom measures, because they measured a specific period of time in the recent past. Precedent for this classification is provided by Todman (2013), who assessed state boredom over the preceding two weeks. All correlations were $p < .001$. MSBS = Multidimensional State Boredom Scale; BPS = Boredom Proneness Scale. The test items did not explicitly indicate state/trait boredom (as written above) in the administered survey. Sample sizes are approximate due to some missing data.

Table 3

Goodness of fit indices for confirmatory factor analysis of full MSBS

	Chi Square	Chi Square df	GFI	CFI	TLI	SRMR	RMSEA	Overall fit
Primary sample, Fahlman et al. (2013) model	3743.035	372	.855	.929	.923	.041	.073 [90% CIs, .071, .075]	Decent-good
Primary sample, arousal items combined	4449.288	373	.825	.914	.907	.042	.080 [90% CIs, .078, .082]	Decent
Primary sample, five first-order factors	3659.363	367	.858	.931	.923	.040	.072 [90% CIs, .070, .074]	Decent-good
Primary sample, one factor only	9733.810	377	.648	.803	.788	.062	.120 [90% CIs, .118, .122]	Poor
Advertisement sample, Fahlman et al. (2013) model	867.085	372	.774	.896	.887	.060	.082 [90% CIs, .075, .089]	Mediocre
Student sample, Fahlman et al. (2013) model	1127.861	372	.834	.905	.896	.062	.069 [90% CIs, .065, .074]	Decent

Notes: GFI: Goodness of Fit Index; CFI: Comparative Fit Index; TLI: Tucker Lewis Index; SRMR: Standard Root Mean Square Residual; RMSEA: Root Mean Square Error of Approximation

Table 4

Goodness of fit indices for confirmatory factor analysis of MSBS-15

	Chi Square	Chi Square df	GFI	CFI	TLI	SRMR	RMSEA	Overall fit
Primary sample, model based on Fahlman et al. (2013)	1043.779	85	.919	.961	.952	.035	.081 [90% CIs, .077, .086]	Good
Primary sample, arousal items combined	1678.289	.86	.861	.936	.922	.041	.104 [90% CIs, .100, .108]	Decent-good
Primary sample, five first-order factors	987.689	80	.924	.963	.952	.033	.081 [90% CIs, .077, .086]	Good
Primary sample, one factor only	4775.598	90	.713	.811	.779	.067	.174 [90% CIs, .170, .178]	Poor
Advertisement sample, model based on Fahlman et al. (2013)	216.656	85	.869	.946	.933	.052	.088 [90% CIs, .074, .103]	Decent
Student sample, model based on Fahlman et al. (2013)	253.911	85	.923	.960	.951	.047	.069 [90% CIs, .059, .079]	Good

Notes: GFI: Goodness of Fit Index; CFI: Comparative Fit Index; TLI: Tucker Lewis Index; SRMR: Standard Root Mean Square Residual; RMSEA: Root Mean Square Error of Approximation

Table 5

Goodness of fit indices for confirmatory factor analysis of MSBS Short Form

	Chi Square	Chi Square df	GFI	CFI	TLI	SRMR	RMSEA	Overall fit
Primary sample, one factor only	465.837	20	.933	.946	.925	.037	.114 [90% CIs, .105, .123]	Decent
Advertisement sample, one factor only	87.511	20	.901	.911	.875	.057	.131 [90% CIs, .103, .159]	Mediocre
Student sample, one factor only	124.838	20	.929	.911	.875	.050	.112 [90% CIs, .093, .131]	Mediocre-decent

Notes: GFI: Goodness of Fit Index; CFI: Comparative Fit Index; TLI: Tucker Lewis Index; SRMR: Standard Root Mean Square Residual; RMSEA: Root Mean Square Error of Approximation

Table 6

Internal consistency (Cronbach's α) for the MSBS (sub)scales and the condensed scales (primary sample)

Subscale	Short Form	MSBS-15	Full Scale	Disenga--gement	High Arousal	Inattention	Low Arousal	Time Perception
Full sample	.909	.961	.976	.944	.923	.898	.929	.928
Males	.914	.963	.977	.946	.920	.895	.926	.921
Females	.904	.959	.976	.942	.926	.900	.931	.931

Note: Analyses included between 670-770 participants for males, 789 - 902 participants for females, and 1462-1679 participants for the full sample. MSBS = Multidimensional State Boredom Scale.

Table 7

Intercorrelations of the MSBS (sub)scales and the condensed scales, and correlations with other boredom and mood measures (primary sample)

Scale	MSBS Short Form	MSBS-15	MSBS Full Scale	Disengagement	High Arousal	Inattention	Low Arousal	Time Perception
MSBS Short Form	1	.927	.962	.964	.823	.864	.834	.747
MSBS-15	.927	1	.988	.929	.907	.827	.944	.773
MSBS Full Scale	.962	.988	1	.964	.903	.849	.916	.787
Disengagement	.964	.929	.964	1	.832	.791	.860	.803
High Arousal	.823	.907	.903	.832	1	.749	.818	.624
Inattention	.864	.827	.849	.791	.749	1	.723	.574
Low Arousal	.834	.944	.916	.860	.818	.723	1	.639
Time Perception	.747	.773	.787	.703	.624	.574	.639	1
ZBS	.280	.270	.281	.285	.261	.203	.243	.229
BPS Total	.794	.802	.815	.804	.696	.688	.742	.656
BPS-AR	.732	.725	.746	.759	.658	.632	.685	.523
BPS-C	.385	.399	.406	.381	.409	.360	.357	.300
BPS-ES	.674	.654	.674	.690	.568	.554	.608	.514
BPS-IS	.455	.488	.484	.454	.411	.432	.447	.416
BPS-PT	.739	.755	.761	.734	.611	.617	.693	.705
SBM item 1	.641	.638	.653	.653	.544	.522	.600	.543
SBM item 2	.575	.591	.600	.586	.508	.470	.558	.515
SBM item 3	.372	.381	.390	.385	.326	.331	.367	.306
SBM item 4	.446	.455	.458	.460	.380	.363	.449	.345
SBM item 5	.639	.668	.667	.654	.568	.536	.650	.517
SBM item 6	.558	.552	.564	.577	.465	.455	.536	.420
SBM item 7	.434	.486	.475	.448	.414	.385	.470	.381
SBM item 8	.556	.584	.585	.574	.511	.481	.578	.419
DASS-D	.719	.796	.782	.748	.717	.631	.807	.537
DASS-A	.553	.616	.605	.558	.571	.522	.595	.446
DASS-S	.635	.701	.694	.643	.731	.600	.665	.447

Notes: *N* ranges from 1679 to 1710, depending on the extent of missing data. All correlations are significant at $p < .001$. MSBS = Multidimensional State Boredom Scale; ZBS = Zuckerman's Boredom Susceptibility Scale; BPS = Boredom Proneness Scale, BPS-AR = Affective Responses subscale, BPS-C = Constraint subscale, BPS-ES = External Stimulation subscale, BPS-IS = Internal Stimulation subscale, BPS-PT = Perception of Time subscale; SBM = State Boredom Measure, DASS = Depression Anxiety Stress Scales, DASS-D = Depression subscale, DASS-A = Anxiety subscale, DASS-S = Stress subscale

Appendix

Exploratory factor analysis of the MSBS Short Form

A series of exploratory factor analyses were run on the MSBS Short Form data from the primary sample (including imputed data).The analyses were run using FACTOR {Lorenzo-Seva, 2013 #993}. A Parallel Analysis based on minimum rank factor analysis was conducted, using the polychoric correlation matrix from the data. Promin rotation was employed. This analysis indicated that there was one dimension/factor in the data. The first three Eigenvalues were 5.102, 0.780, and 0.656. Only the first Eigenvalue was significantly greater ($p <$ .05) than the corresponding randomly generated Eigenvalue. A second analysis was run on the data using the same specifications, except with a Minimum Average Partial Test rather than a Parallel Analysis. This test also indicated that the data contained a single dimension/factor.