IDENTIFYING COVER PAGE

TITLE

Is Mindfulness a Noticeable Quality? Development and Validation of the Observed Mindfulness Measure

AUTHORS

Larissa Bartlett [1, 2] ORCID: 0000-0003-0474-8107

Angela J Martin [1, 3] ORCID: 0000-0003-0109-1218

Raimondo Bruno [4] ORCID: 0000-0001-6673-833X

Michelle Kilpatrick [1] ORCID: 0000-0001-6039-9154

Kristy Sanderson [1, 5] ORCID: 0000-0002-3132-2745

Amanda L Neil [1] ORCID: 0000-0002-1344-6672

AFFILIATIONS

1. Menzies Institute for Medical Research, University of Tasmania
2. Wicking Dementia Research and Education Centre, University of Tasmania
3. Pracademia Pty Ltd
4. School of Medicine, University of Tasmania
5. Faculty of Medicine and Health Sciences, University of East Anglia

CORRESPONDING AUTHORS

Dr Larissa Bartlett

Wicking Dementia Research and Education Centre

University of Tasmania

Private Bag 143

Hobart 7001

Tasmania Australia

Larissa.Bartlett@utas.edu.au

Tel: +61 3 6226 4723

Dr Amanda Neil

Menzies Institute for Medical Research

University of Tasmania

Private Bag 23

Hobart 7001

Tasmania Australia

Amanda.Neil@utas.edu.au

Tel: +61 3 6226 4640

KEYWORDS

Mindfulness, Social behaviour, Measurement, Observer-reports

ACKNOWLEDGEMENTS

This project was conducted at the Menzies Institute for Medical Research.

The project was conceived and progressed by LB as part of her PhD (Medical Sciences) at the Menzies Institute for Medical Research, with supervision by AN (Primary), KS, AM and MK. LB has since commenced a post-doctoral appointment with the Wicking Dementia Research and Education Centre. The study team gratefully acknowledges Dr Bruno Cayoun, Professor Ruth A Baer, Professor Harald Walach, Dr Craig Hassed, Pamela Lovell and Professor Gary Johns who acted as Expert Advisors, Tim Albion for technical assistance, and the study participants. The results of this study have been presented in poster format at the 2018 European Academy of Occupational Health Psychology Conference and as an oral presentation

at the 2018 Mind and Life Institute's International Symposium on Contemplative Research.

FUNDING

LB's PhD studies were supported by an Australian Government Research Training Program PhD Scholarship, supplemented by the TasNetworks Mental Health and Wellbeing Elite PhD Scholarship.

CONFLICTS OF INTEREST

The authors declare no potential conflicts of interest.

ETHICS APPROVAL

This study was undertaken in accordance with the ethical standards of the Tasmania Social Sciences Human Research Ethics Committee (ref: H0016587). Informed consent was obtained from all of the participants in the study.

AVAILABILITY OF DATA

Study data will be made available by the corresponding author on request.

ABSTRACT

Objective: This paper presents the psychometric development of a new observer-report research questionnaire for assessing aspects of an individual's mindfulness that are noticeable to others.

Methods: Items from five established self-report mindfulness scales were re-worded for observer-report, and 30 were endorsed for potential inclusion by an expert panel (n=5). Factor analytic and item response theory models were used to test item and scale psychometrics with data (N=494) in three independent samples. The resultant scale was assessed for reliability, validity and responsiveness.

Results: A nine-item, three-factor scale with good fit indices was determined (RMSEA=0.04, CFI=0.99, TLI=0.99). The instrument provides an overall score for observed mindfulness and subscale scores for observed attentiveness, awareness and acceptance. Within-subject test-retest reliability was strong (ICC 0.91) and agreement between observed and self-reported mindfulness was adequate (ICC 0.45). Validity tests showed concordance between the new scale and self-reported mindfulness within the selected nomological network constructs (emotional intelligence, empathy and avoidant behaviours). Responsiveness was indicated but unconfirmed in data from a randomized controlled trial of low-dose mindfulness training.

Conclusion: The Observed Mindfulness Measure (OMM) is a quantitative instrument that can provide an additional data source to strengthen self-reported findings in mindfulness research. With some further refinement this new instrument can advance research into whether and how mindfulness training might make a difference in social and organizational domains.

KEYWORDS

Mindfulness, Social behaviour, Measurement, Observer-reports

Is mindfulness a noticeable quality? Development and validation of the Observed Mindfulness Measure

Mindfulness is characterized by intentionally paying attention to current experiences with an open and curious attitude (Bergomi et al., 2013; Bishop et al., 2004). This natural human quality is consistently associated with positive psychological resources including resilience, optimism, self-efficacy, wellbeing and life satisfaction (Jha et al., 2010; Khoury et al., 2015; Roche et al., 2014).[1] There is a growing body of work that suggests the benefits of a persons' mindfulness may manifest *beyond* individual outcomes into social and behavioural domains (Donald et al., 2019; Singh et al., 2015; Van Doesum et al., 2018; Van Lange and Van Doesum, 2015).

The premise that mindfulness may influence social interactions is in keeping with the contemplative teachings from which the construct and practices have emerged into contemporary western culture. In the Sutta, the 'backbone' of Buddhist teachings (Thera, 1998), foundational mindfulness practices encompass paying intentional, non-judgmental attention to internal (self) and external (other) experiences (Supplementary Figure S.1.). Through the sustained practice and application of mindfulness one is likely to be aware not just of the impact of one's own experiences, but that awareness of others' experiences can also induce emotional reactions and behavioural urges. Van Doesum's (2013) theory of social mindfulness articulates the relationship of one's mindfulness with empathy (particularly perspective taking and empathic concern), wisdom (intentionally seeing views of others) and moral conduct (acting to benefit others). Thus, mindfulness may assist in cultivating both the will and the skill to navigate social situations and relationships (Van Doesum et al., 2013).

[1] Readers are invited to view the Online Supplementary Materials for a summary description of the mechanisms and outcomes of mindfulness practice as they cultivate internal and external experiences.

Prosocial behaviour involves acting voluntarily to maximise the welfare of others (Eisenberg, 2015) and is enhanced by skills such as meta-awareness, decentering and reperceiving, which are cultivated during mindfulness training (Block-Lerner et al., 2007; Eisenberg, 2015). Meta-analytic evidence (N=13,820; Donald et al., 2019) shows that mindfulness and prosociality are moderately and positively correlated. Mindfulness also correlates with emotional intelligence, empathy and non-attachment behaviours, which are known determinants of prosocial behaviours (Miao et al., 2018; Jaurequi, 2019; Donald et al., 2019; Davis, 1983; Cardaciotto et al., 2008; Baer et al., 2004; Block-Lerner et al., 2007; Aydin Sünbül, 2019). Further, Karremans et al. (2017) argues that pro-relationship motivations and behaviours such as sacrifice, forgiveness, partner acceptance and limiting stress spill-over are influenced by capacities for emotional regulation and executive control, which are enhanced through increased mindfulness (Chambers et al., 2009; Malinowski, 2013). Thus, there is growing theoretical and evidentiary support for mindfulness as a noticeable quality that is instrumental in supporting prosocial behaviours. In practice, skilful social interactions are central to effective clinical care, classroom learning and workplace dynamics, and mindfulness training may offer a means for improving service quality and client outcomes in these contexts.

Assessing mindfulness as a quality noticeable to others

There are several self-report mindfulness measures used in research (Baer et al., 2009; Qu et al., 2015), but despite there being interest in its potential social effects there are no quantitative measures currently available that assess the degree to which another person behaves mindfully. Existing research investigating interpersonal outcomes of mindfulness has used self-report mindfulness questionnaires (completed by a person about themselves) in conjunction with qualitative or count data provided by known observers (e.g. family, friend, patient, colleague of that person). Using this approach a number of studies have provided evidence that a person's mindfulness is noticeable and potentially beneficial within the context

of work (Good et al., 2016; Hyland et al., 2015; Reb and Choi, 2014), intimate relationships (Karremans et al., 2017; Khaddouma et al., 2017; Gillespie et al., 2015; Barnes et al., 2007), child aggression and parental discipline (Singh et al., 2010; Siu et al., 2016). Positive associations have also been found in health care settings between clinician mindfulness and patient-reported quality of care (Horner et al., 2014; Cohen-Katz et al., 2005; Singh et al., 2015). Further, in education settings, school teachers have reported being more able to reappraise challenging situations following mindfulness training, leading to improved self-care and student engagement (Sharp and Jennings, 2016; Taylor et al., 2016).

Valid and reliable self-report questionnaires provide an important and quantifiable perspective on the effects of behavioural interventions (DeVellis, 2012; Varela and Shear, 1999). However, reliance on single source, self-reported data has been identified as a methodological limitation in mindfulness research (Allen et al., 2015; Choi and Leroy, 2015; Davidson and Kaszniak, 2015; Goldberg et al., 2017; Grossman, 2018; Van Dam et al., 2017). Specifically, self-report assessments can be subject to response bias through expectancy effects and social desirability (Kreplin et al., 2018; Vazire and Mehl, 2008; Rosenthal, 2009; Van de Mortel, 2008), and in the absence of triangulating evidence, findings may conflate the respondents' inner states, motivations and influences with reflections of behaviours or predispositions (Baumeister et al., 2007). These concerns affect half the published mindfulness evidence, which is based solely on self-report assessments (Goldberg et al., 2017).

Using additional data sources to cross-validate, or triangulate, findings from self-report data would increase confidence and strengthen research in the field (Goldberg et al., 2019; Davidson and Kaszniak, 2015; Hyland et al., 2015). Objective measurements and qualitative data have been used to triangulate self-reported findings (Van Dam et al., 2017). Objective measurements include physiological markers (e.g. neurological imaging, hormonal fluctuations, blood pressure and heart rate variability) (Chiesa and Serretti, 2010; Garland et

al., 2017a) and cognitive performance tests (e.g. attentional control) (Chiesa et al., 2011; Creswell and Lindsay, 2014). However, collecting these objective data is costly even for small studies, and logistically prohibitive for large cohort studies or field trials. Qualitative data are less expensive to collect than objective markers and are well qualified for providing a rich source of information about study participants in behavioural research (Madill and Gough, 2008). For example, open-ended participant reports have contributed to current phenomenological understanding of the effects and experiences associated with mindfulness training and practice (e.g. Eby et al., 2017; Cohen-Katz et al., 2005; Gillespie et al., 2015; Hugh-Jones et al., 2018). Summary statistics from this type of qualitative data have been used to support self-reported findings in a small number of studies (e.g. Bartlett et al. 2016) but the variety of paradigms and methods used in qualitative research limits the potential for synthesizing findings (Lewin et al., 2015; Cairns and Murray, 2015; Sandelowski et al., 1997). In contrast, observations collected using a quantitative second-person report measure could be pooled through meta-analysis to enable a synthesis of evidence from multiple studies.

There are currently two self-report measures designed specifically to assess one's own interpersonal mindfulness; the Interpersonal Mindfulness Scale (IMS) (Pratscher et al., 2018) and the Interpersonal Mindfulness in Parenting Scale (IM-P) (Duncan, 2007). Both these multi-faceted instruments contribute to researcher understanding about the aspects of individual mindfulness and how it might impact on others. However, these are both self-report scales and neither quantifies the extent to which mindfulness manifests in behaviours noticeable to others and thus may exert effects beyond the individual practitioner. A review of available published literature revealed just one study has attempted to validate an observer administered assessment protocol, the Child Observation of Mindfulness Measure (C-OMM) (Lemberger-Truelove et al., 2019). The C-OMM was designed for detecting degrees of mindfulness in young children (aged 3 to 4 years of age). The administration of the C-OMM involves multiple raters

conducting multiple assessments and includes considerations specific to research related to young children. This work illustrates the interest in, and potential value of, observer assessments in mindfulness research.

An alternative to direct observation is the use of a research questionnaire that collects quantitative second-person reports. Second person reports are provided by a known other (e.g. family, friend or colleague). They differ from first person, or self-report data that reflects what the subject may say about him or herself, and third person assessments that are made by someone with no prior relationship with the subject (Varela and Shear, 1999). Second-person reports provide an outside-view or perspective on the behaviour of a paired self-reporter (Varela and Shear, 1999). While not free of procedural and contextual bias, a validated quantitative instrument that can collect observations from known others may provide useful data for cross-validation, and information about study subjects that may not be available from the first-person, self-report perspective (Atkins and Wood, 2002; Baumeister et al., 2007; Floyd, 2005; Markon et al., 2013; Olino and Klein, 2015; Vazire and Mehl, 2008). Second-person reports can be compared directly with self-report data to understand the extent to which the two perspectives converge (known as self-other agreement, or SOA) (Funder and West, 1993). Frequently used in occupational, psychological and educational research and practice, second-person reports are easy and inexpensive to collect (Vazire, 2006).

Two studies have shown that collecting quantifiable, second-person reports of another person's mindfulness using research questionnaires is feasible and informative (Bartlett et al., 2016; May and Reinhardt, 2018). These studies asked participants to answer some questions about themselves (self-report) and to nominate a person well known to them to act as observers, to answer questions about their paired participant's behaviours (observer-report). The two studies compared the degree of agreement between self-report data collected from established

self-report mindfulness measures and observer-report data collected from adapted versions of the self-report questionnaires.

The 15-item Mindful Attention and Awareness Scale (MAAS; Brown and Ryan, 2003) was used by Bartlett et al. (2016), to measure changes in participants' mindfulness following a five-week mindfulness course in a sample of Australian public sector employees. Participants were invited to nominate up to two trusted friends, family members or work-based contacts with whom they interact two to three times a week to act as their observer(s). Observers were invited to answer some questions about their paired participant in surveys administered online, at the same time as the participant surveys (baseline and post-intervention). For example, *I break or spill things because of carelessness, not paying attention, or thinking of something else* was changed to *The person tends to have a lot of minor mishaps or accidents*. Responses were recorded (1 to 6) and mean scores calculated in alignment with the MAAS method. There was negligible agreement between the observer-reported and self-reported change in mindfulness in this study (r=0.06, ns). However, the observers' free-text reports indicated that compared with the inactive control group, the course participants were noticeably less emotionally reactive and stressed, and more attentive, vigorous and productive (Bartlett et al., 2016). The authors conclude their adaptation of the MAAS items for observer report did not detect the changes that were noticed by observers of course participants, and that a validated measure that has good construct validity would be more informative. May and Reinhardt (2018) investigated agreement between self- and observer-reports using the 39-item Five Facet Mindfulness Questionnaire (FFMQ; Baer et al., 2006). Questions were adapted for observer-report by switching the personal pronouns ('I') to second-person pronouns ('he' or 'she') and scoring followed established guidelines. This study found consistent low-level agreement between self- and observer-report data. Mindfulness facets that scored higher for observability (i.e. were outcome focused, rather than process focused) achieved higher self-other agreement.

May and Reinhardt (2018) concluded that the internal processes of mindfulness may manifest in overt behaviours, which are easier to notice than internal cognitive and emotional processes (Funder and Dobroth, 1987). Understanding the extent to which a person's mindfulness influences behavioural outcomes is thus best advanced by retaining a focus on aspects of mindful acting that are higher in observability.

Self-other agreement means the degree to which our own perception of ourselves concords with how others see us (Funder and West, 1993). If self-other agreement is high, a stronger level of confidence in self-reported information can be inferred. Inconsistent or low levels of self-other agreement can be due to disagreement, where there is a real difference in perceptions of the two reporters about the assessed attribute(s). Low agreement may, however, be attributable to method variance, where the responses are biased due to differences in the information available to the reporter, or differences in the perceived desirability of traits being assessed (Vazire and Mehl, 2008). Further, the data collection instruments used to assess the two perspectives may not tap into the same constructs, or the constructs being assessed are not as noticeable to one of the parties (e.g. internal processes vs outcomes/behaviours). The studies conducted by May and Reinhardt (2018) and Bartlett et al. (2016) suggest that while an individual's mindfulness is essentially an internal process or trait, it does manifest in behaviours that are noticeable to others. Findings regarding the extent to which a person behaves in a way that is noticeably mindful may be more robust and defensible if data are collected using a purposefully designed and psychometrically validated scale.

Aims and hypotheses

The ultimate objective of this project was to develop a valid, scalable and inexpensive tool for cross-validating self-reported findings and advancing research into the qualities of mindfulness that are noticeable to others. We first aimed to develop a psychometrically sound second-person report research instrument, the Observed Mindfulness Measure (OMM) to

assess the presence and extent to which a known other behaves mindfully (Aim 1). Our second aim was to test the validity, reliability and responsiveness of the new measure (Aim 2).

Self-reported mindfulness was the key criterion against which the measure was tested. Criterion validity relates to the empirical association of a scale with a pre-determined criterion or 'gold standard' where one exists, whereas construct validity relates to the associations of the scale with validated measures in a nomological network of related constructs (Cronbach and Meehl, 1955; DeVellis, 2012). Criterion and construct validity for a second-person report mindfulness measure are indicated if participants' self-reported mindfulness converges with observers' reports about the participants' mindful behaviours. It was anticipated that scores on the OMM would show convergence with self-reported mindful attention and awareness (MAAS; Brown and Ryan, 2003), which would suggest criterion validity. However, we expected the degree of self-other agreement would be subject to method variance because of the differing perspectives of the respondents, leading to a lower correlation than would be observed between two convergent self-report measures. Our criterion validity hypothesis was therefore that there would be a small to moderate sized positive empirical association between individual's self-reported mindfulness (MAAS) and second-person observations of that person's mindful behaviours (OMM) (H1a).

To determine construct validity, we tested the correlations between data collected using the second-person reported OMM and self-report measures of constructs in the nomological network of mindfulness. Only those constructs which are understood to influence social interactions were included: emotional intelligence, empathy and non-attachment. Prior work has shown emotional intelligence has a consistent, moderate to strong association with self-reported mindfulness ($\bar{r}$=0.41) (Miao et al., 2018). We hypothesized the OMM would show a small positive association with the clarity of feelings and mood repair dimensions within the

Trait Meta Mood Scale (TMMS; Salovey et al., 1995), and a small negative association with the attention to feelings dimension (H1b).

Empathy can be cultivated through mindfulness training (r=0.44) (Kreplin et al., 2018; Block-Lerner et al., 2007; Van Doesum et al., 2013) and emotional distress is moderately and negatively associated with mindfulness (Hedge's g=0.68) (Bartlett et al., 2019; Virgili, 2015). We therefore hypothesized that, using the Interpersonal Reactivity Index (IRI; Davis, 1983), the empathic concern and perspective taking dimensions of empathy would be convergent with OMM data, and show a small to moderate positive correlation, while emotional distress was expected to return a similar sized negative correlation as a divergent construct (H1c).

The Acceptance and Action Questionnaire (AAQ; Hayes et al., 2006) assesses avoidant behaviours which have been shown to be negatively associated with mindfulness (r= -0.30 to r= -0.51) (Baer et al., 2004; Cardaciotto et al., 2008). We therefore hypothesized the AAQ and OMM scores would have a small to moderate sized negative correlation (H1d).

To assess reliability, we hypothesized the measure would be stable in test-retest, with a intra-rater reliability score of 0.8 or higher, which would indicate temporal stability (H2a). And lastly, to test responsiveness of the scale to change, we hypothesized that OMM data gathered prior to and after a mindfulness training program would show changes in the same direction, but lower in magnitude to changes observed in self-report MAAS data (H2b).

METHODS

Participants

Three independent samples provided data used for this project. Table 1 presents the characteristics of each sample at baseline. Eligibility criteria for samples A and B required participants to be 18 years of age or over. Sample C were deemed eligible for the primary study they were recruited into if they were over 18 years of age, and able to participate in the

intervention and met mental health screening criteria for an unsupported meditation course (Bartlett et al., 2021).

Sample A was a community sample of adults invited via a social media campaign to help develop a new measure for mindfulness research. This sample of individuals responded to a study advertisement with the title "Can you tell how mindful someone else is?" The study information and consent form were presented if the advertisement link was followed. Respondents were asked to think about someone they see frequently (two to three times a week) and know well, and to rate that person's behaviour in response to a series of statements. The resultant sample A (n=307) was mostly female (76%), aged between 35 and 65 years with 92% reporting some prior mindfulness exposure. This sample provided responses to the list of potentially included items, with their observations referring to work colleagues (14%), friends (28%) and family members (58%). The data provided by Sample A were used to test the list of potentially included items for interpretability, to explore the structure (or factors) in the data, and to test item performance.

Sample B were participant-observer dyads recruited via a separate social media campaign with the title "Is mindfulness a noticeable quality?", calling for people with an interest in joining a mindfulness research project. After following the link in the advertisement, interested respondents ('participants') were presented with study information and consent. Before progressing to their surveys, these participants were asked to nominate a well-known and trusted colleague, family member or friend to be paired with them for the study ('observers'), and who would be willing to answer some questions about the participant's behaviours. If an observer was nominated, an invitation was generated to the nominee advising that [participant's name] had invited them into the study. Participant-observer dyads were formed if the observers responded to the invitation and enrolled in the study. Cases were included in analyses if both members of Sample B dyads provided full data. The resultant sample B

included 88 dyads comprised of 176 individuals, mostly located in Australia (93%) and evenly distributed by sex and age. This sample provided observer reports for work colleagues (11%), friends (27%) and family members (61%) and self-report data from the people being observed (participants). Sample B data were used in tests of temporal stability, confirmatory factor analyses, item performance testing and in conjunction with self-report data, to test the construct validity of the new measure.

Sample C was also comprised of dyads, and was drawn from a randomised controlled trial (RCT) of workplace-delivered app-based mindfulness training (ACTRNXXXXXXX) (Bartlett et al., 2021). Recruitment into the RCT involved inviting program participants to nominate a work-based observer using the same method as for Sample B. Both members of Sample C participant-observer dyads answered survey questions before (baseline) and after training (post-intervention) and then six-months from baseline (follow-up). Sample C included 99 participant-observer dyads comprised of 198 individuals who were mostly female (participants 79%, observers 74%) and aged between 35 and 64 years. This sample were all Australian public sector employees. The observers in Sample C dyads included co-workers (21%), subordinates (14%), supervisors (17%) or work-based friends (43%) of their paired participants. Sample C data were used to support factor analyses, item performance testing and to assess the responsiveness of the new measure to change.

< Insert Table 1 about here >

Procedures

Construct definition

Observed mindfulness is conceptually a *new view of an established construct,* so the commonly accepted two-component definition of (personal) mindfulness put forward by Bishop et al. (2004) was adapted. The working definition of observed mindfulness employed

in the development of the OMM is *the noticeable tendency of another person to be mindful: attentive to and aware of current experience, and displaying an attitude of curiosity, openness and acceptance.* Observed mindfulness therefore does not refer to the internal experiences of the subject being observed, rather it focuses on the degree to which the subject noticeably acts or responds mindfully in social contexts.

Item selection

A list of 102 potential items was created by pooling items from five established self-report instruments: the Mindful Attention and Awareness Scale (MAAS) (Brown and Ryan, 2003); the Freiberg Mindfulness Inventory (FMI) (Walach et al., 2006); the Kentucky Inventory of Mindfulness Skills (KIMS) (Baer et al., 2004); the Philadelphia Mindfulness Scale (PHLMS) (Cardaciotto et al., 2008); and the Langer Mindfulness Scale (LMS) (Pirson et al., 2018). These measures are some of the most frequently used instruments in mindfulness research and collectively represent the range of construct conceptualizations (Qu et al., 2015; Bergomi et al., 2013). All items were re-written in observer-report format (A1, A2, A4) with advice from a Visiting Scholar (Anon) with expertise in scale development. The directionality of each item (positive or negative wording) was retained in the translation and items that did not preserve the same conceptual meaning were discarded. Of the original 102 items, 51 were considered to have retained face validity and were sent to a panel of expert advisors for review. Advisors (n=16) with a range of expertise and perspectives were identified by reviewing mindfulness training and research literature and were invited by email to contribute to the study. Five advisors provided considered responses, including lead authors of previously validated self-report measures of mindfulness (n=2), established mindfulness researchers with an interest in assessment (n=1), and/or certified and experienced mindfulness teachers (n=2). Advisors rated each translated item as acceptable, maybe acceptable or not acceptable, in terms of whether the original meaning of the self-report item was retained, and if they considered the

translated item potentially appropriate for detecting observed mindfulness. Advisors provided brief explanations for their decisions. Items endorsed by at least three advisors were considered to have face validity and were included in the item pool for interpretability and psychometric testing.

The list of 51 items worded for observer report was reviewed by five expert advisors who provided advice on the items' face validity (Supplementary Information Table S.1). Thirty items were endorsed for retention by at least three of the expert advisors, the remaining items were dropped.

Survey administration

Samples A, B and C provided data via online surveys that were developed and administered using Research Electronic Data Capture (REDCap) (Harris et al., 2009) hosted at the University of XXXX. Individuals' responses across time points and within dyads were linked by code and all analyses were conducted on de-identified data. The survey administration schedule for each respondent sample is available in Supplementary Information Table S.2. All items endorsed by the expert advisors were presented to Sample A. Samples B and C were presented with reduced item sets as determined by the development process.

Potential OMM items were presented to observers in all samples with response options on a five-point Likert scale. This format was chosen because it is used by four of the five measures from which the items were drawn. Respondents were asked to "Please tell us how frequently each of the following statements would be true about your paired participant". Response labels were 1: "Not at all", 2: "Occasionally", 3: "About half the time", 4: "Frequently" and 5: "All the time".

Interpretability

Sample A data was used to assess the interpretability of each of the items retained after expert review. The items were presented with the five-point Likert scale as described above with an extra option (0) labelled "Cannot answer". When "0" was selected respondents were asked to give a brief explanation. Based on guidelines recommending the retention only of items that are interpretable for the majority of participants, items were dropped if 2% respondents selected "0" (DeVellis, 2012). Only one item ("The person makes judgments about how worthy his or her experiences are") was considered difficult to answer by participants (n=11, 4%). This item was excluded, and factor analyses commenced with 29 items.

Measures

Demographics and dyadic relationship information were collected at baseline in all samples. The following self-report research instruments were administered in addition to the observer-reported OMM items, to collect data for testing the agreement between observer-report and self-report data, and to understand the construct validity of the new measure.

The Mindful Attention and Awareness Scale (MAAS) (Brown and Ryan, 2003) is a well-established 15-item self-report measure of mindfulness that has been used as a reference construct in other mindfulness measure development studies (Pratscher et al., 2018; Cardaciotto et al., 2008; Pirson et al., 2018). The Trait Meta Mood Scale (TMMS) (Salovey et al., 1995) assesses emotional intelligence. Its three subscales include attention to one's feelings, clarity of these feelings, and mood repair or ability to regulate one's feelings. The sub-scales are correlated, with a negative association between attention to feelings and the other two. The Interpersonal Reactivity Index (IRI) (Davis, 1983) empathy scale assesses empathic concern, fantasy, perspective taking and emotional distress. The Acceptance and Action Questionnaire

(AAQ) (Hayes et al., 2006) is a nine-item measure of experiential avoidance, defined as negative evaluation of, and unwillingness to maintain contact with, internal experiences such as bodily sensations, emotions, cognitions, and urges.

Statistical methods

Analyses were conducted using the R statistical computing environment (R Core Team, 2014). The Psych (Revelle, 2019), GPARotation (Bernaards and Jennrich, 2005), lavaan (Rosseel, 2012), mirt (Chalmers, 2012) and ltm (Rizopoulos, 2006) packages were used, with guidance from Revelle's Personality Project online resource (Revelle and French, 2013) and the step-by-step guide to Latent Variable Modeling Using R by Alexander Beaujean (2014). Incomplete cases were excluded from analyses because there were a sufficient number of complete responses and insufficient information was available to help identify patterns of missingness and direct imputation decisions (Dong and Peng, 2013; Beaujean, 2014). Significance tests were conducted with alpha set to 0.05. Prior to analyses, negatively worded items were reverse-scored to ensure consistent directionality and all variables were checked for normal distribution using Q-Q plots.

Structure and item retention

Factor analyses were applied with the maximum likelihood estimator and oblique rotation to determine the structure of the data, using a polychoric correlation matrix as appropriate for ordinal polytomous data. Model fit indices reported are the Comparative Fit Index (CFI) and Tucker Lewis Index (TLI), which indicate good fit if scores are greater than 0.95; the Root Mean Square Error of Approximation (RMSEA), which indicates poor fit if the score is higher than 0.05; and χ^2 values, which indicate the model fits the data reasonably well if not significant (Beaujean, 2014). Exploratory factor analyses (EFA) were run on Sample A responses to the 29-item set to test the factorial structure of the measure. The fit for one through

five dimensions were compared to determine the best model. Items with a factor loading of at least 0.5 and a minimum difference of 0.3 from the next strongest factor loading in the determined model were retained for further testing. These criteria were set to ensure each item in the scale was clearly associated with just one of the factors (Smith et al., 2003; DeVellis, 2012). Observers' baseline responses from pooled independent Samples B and C were then used for confirmatory factor analyses (CFA) on the reduced item set.

Item responsiveness tests (IRT) were conducted using pooled observer baseline data from Samples A, B and C. Models were fitted with a General Partial Credit Model for ordinal polytomous data (Muraki and Muraki, 2016). This procedure uses generalised linear models to identify differential item functioning, and the likelihood of respondents with a certain level of the assessed trait endorsing a particular response option (Hays et al., 2000). Discrimination scores (*a*-values) indicate how well the item differentiates those higher or lower on the trait and should be close to or higher than one. Item characteristic trace plots illustrate the amount of information obtained for each response option. Difficulty is indicated by the threshold scores (*b*-values), where the probability of a certain response is calculated as a function of the difference between the respondent's trait level and the intersection parameter. Low and high scores indicate lower and higher levels of the measured trait respectively. Information curves with the greatest area below zero on the x axis suggest that respondents with a below average score on the measured trait will be well identified by the item. Decisions to flag items for potential exclusion were based on both semantic review and IRT statistics. Items that scored poorly for precision, information or were potentially redundant were flagged for possible removal. Systematic exclusion of these flagged items then followed in a series of CFAs using pooled data from Samples B and C to identify the final scale.

McDonald's Omega test of factor saturation was run on the final model to check item-factor relationships and to confirm the presence of an over-arching general factor (Revelle and

French, 2013). The Omega test employs the parallel analysis estimator to compare real and random data and then applies a Schmid Leiman transformation to obtain an index for hierarchical (ω_h) and total (ω_t) factor saturation. The Omega coefficients are considered a more stable index of internal consistency for multidimensional models than Cronbach's alpha (α) and Guttman's Lambda (λ) (also reported) (Dunn et al., 2014). IRT was re-run on unidimensional data for each factor in the determined model. Table 2 shows the step-wise item reduction decision making process.

< Insert Table 2 about here>

Validity

To assess the new scale for criterion validity (H1a) the degree of self-other agreement between scores from the observer-reported OMM and self-reported MAAS was tested, followed by the agreement between OMM scores and participants' responses to the original nine OMM items in self-report format. We applied tests of *inter*-rater agreement using a one-way, random effects model (Connelly and Ones, 2010). Construct validity was indicated if dyadic data showed convergence, divergence and independence between the OMM and constructs in the nomological network (H1b, H1c, H1d). Pearson's correlations were applied to test the associations of OMM scores at scale and factor level with the computed scores from the self-report measures described above. Based on the degree of self-other agreement in established instruments (e.g. 306 Degree Assessment Tool; Atkins and Wood, 2002), a convergent association of 0.25, or divergent association of -0.25 supported construct validity.

Stability and responsiveness

Sample B completed two surveys, conducted one week apart to test the degree to which OMM responses were consistent when no change was expected (H2a). To assess temporal stability of the scale, intraclass correlation coefficients with a two-way, *intra*-rater agreement

model set for mixed effects were applied to individual data from Sample B. Responsiveness of the OMM (H2b) was tested using complete case data from the experimental and control arms of a randomised controlled trial of an app-based mindfulness program (Sample C) (Bartlett et al., 2021). Observer-reported mindfulness (OMM) and self-reported mindfulness (MAAS) were collected pre- and post-intervention to see whether the OMM was responsive to changes that were present in self-reported data following mindfulness training. Change was first visually inspected by plotting raw change scores by group. Variance in OMM data was assessed using Levene's test of equal variance, and Shapiro Wilk's and Kolmogorov Smironov tests were used to examine data distribution. The difference in mean change at post-intervention for two independent groups was then assessed using independent sample t-tests. Standardized mean difference effect estimates (Cohen's *d*) computed from group mean change data are reported for interpretability (Fu et al., 2013; Kirkwood and Sterne, 2013). Analyses of follow-up data used repeated measures linear mixed models.

RESULTS

Scale development

No clear patterns were observed in the raw correlation matrix of Sample A responses (n=307) to the 29 potentially included items (data not shown). Unrestrained parallel analysis suggested a multidimensional model would fit the data, with up to five components indicated. Fit indices from maximum likelihood EFA were best for a three-dimensional structure (RMSEA=0.083, TLI=0.775). A unidimensional model also returned the same number of items loading as the three-factor model, but the fit was poorer (RMSEA=0.115, TLI 0.565). The two, four and five factor models were ruled out due to cross-loading of items on multiple factors, or insufficient (one or two) items loading per factor. Item content and loadings for the single and three-dimensional models in the 29-item set are presented in Table S.3 (Supplementary

Information). Items achieving cut points were retained, resulting in an 18-item set for further testing. Factor labels, awareness, acceptance and attentiveness, were derived based on semantic review of item content.

< Insert Table 3 about here >

Confirmatory factor analyses confirmed the three-factor structure for the retained 18-items using pooled independent Samples B and C (N=187). IRT modelling identified twelve items with low information or discrimination scores, poor factor loading or potential redundancy, and these were flagged for possible removal (Table 3). Systematic exclusion of these flagged items followed in 22 iterations of CFA using pooled Samples B and C (Supplementary Table S.4). The fit indices for the final model are presented in Table 4. A nine-item scale was determined, with three factors and strong fit indices (RMSEA=0.019, χ^2 p=0.364, CFI=0.998, TLI=0.997). Including Sample A with Samples B and C returned similarly strong results (RMSEA=0.04, χ^2 p=0.01, CFI=0.99, TLI=0.99) as did separate sample analyses (Sample A: RMSEA=0.04, χ^2 p=0.03, CFI=0.99, TLI=0.98; Sample B: RMSEA=0.047, χ^2 p=0.23, CFI=0.984, TLI=0.976; and Sample C: RMSEA=0.053, χ^2 p=0.11, CFI=0.984, TLI=0.977). Correlations between factors were strong (awareness:attentiveness r=0.4, attentiveness: acceptance r=0.5, awareness:acceptance r=0.5). The potential for item reduction to have collapsed the factorial structure was tested by running factor analyses on the unidimensional model, but the three-factor solution prevailed.

< Insert Table 4 about here >

The nine included items all showed trace plots with clear points at which one of the five response options would be most likely, and the information curves suggested the measure detects people at different levels of the underlying trait (Figures 1 and 2). McDonald's Omega test of factor saturation confirmed three inter-correlated factors with an over-arching common

latent variable (Figure 3). Reliability coefficients were adequate for the whole scale (α=0.85, λ_6=0.88, ω_h=0.70, ω_t=0.91) and its sub-scales: awareness (α=0.87, λ_6=0.82, ω_h=0.67, ω_t=0.87), acceptance (α=0.82, λ_6=0.75, ω_h=0.34, ω_t=0.75) and attentiveness (α=0.87, λ_6=0.82, ω_h=0.67, ω_t=0.87). The content of retained items accorded with the definition of observed mindfulness and factor labels awareness, acceptance and attentiveness. As all items included in the final model were clearly associated with only one factor, mean scores could be computed, however summed scores will provide the most information at whole scale and factor level (DiStefano et al., 2009). Thus, the summed total and factor scores from the nine-item OMM were used to test reliability, validity and responsiveness.

<Figure 1 Title: Item characteristic curves for the nine items in the Observed Mindfulness Measure

Figure 1 Legend: Item trace curves for items within each factor from IRT tests using generalised partial credit modelling for ordinal, polytomous data.>

<Figure 2 Title: Information curves for the three dimensions of the Observed Mindfulness Measure

Figure 2 Legend: Factor-level information curves from IRT tests using generalised partial credit modelling for ordinal, polytomous data.>

<Figure 3 Title: Omega tree diagram: factor and scale saturation.

Figure 3 Legend: McDonald's Omega test diagram depicting item-factor and item scale correlations. g: common, over-arching latent trait assessed by the whole scale. F1, F2 and F3 latent trait assessed at factor level.>

Temporal stability

No respondents reported commencing mindfulness training, attending a mindfulness retreat or workshop, or experiencing a particularly stressful event between the two surveys, so all data (n=88 dyads) were included in analysis. Individual observers' responses to the nine-item OMM remained stable at both time points (ICC 0.91, 95%CI 0.85, 0.94). Spearman's monotonic correlations were strong at whole scale (r_s=0.85) and factor levels (Awareness r_s=0.85; Attentiveness r_s=0.80; and Acceptance r_s=0.88). These data suggest excellent test-retest reliability (Koo and Li, 2016).

Criterion and construct validity

Construct validity was tested in Sample B (n=88 dyads). Results of criterion tests showed a strong positive association between observers' total OMM scores and their paired participants' MAAS scores (ICC 0.45, p<0.01, 95%CI 0.16, 0.64). Dyadic (self-other) agreement for the OMM items worded for self- as well as observer-report also showed concordance at whole scale (ICC 0.50, p<0.01 , 95%CI 0.24, 0.67) and factor level (Awareness: ICC 0.26, p=0.08, 95%CI -0.13, 0.51, Attentiveness: ICC 0.36, p=0.02 , 95%CI 0.03, 0.58 and Acceptance: ICC 0.51, p<0.01, 95%CI 0.25, 0.68).

The correlations between the total OMM and dimension scores with the nomological net constructs used in this study are presented in a matrix (Table 5) and illustrated in Supplementary Figure S.2. As hypothesised, OMM scores were convergent with self-reported mindfulness (H1a, MAAS, r=0.29, p=0.01), two dimensions of the self-report measure of emotional intelligence (H1b, TMMS): clarity of feelings (r=0.27, p=0.01) and mood repair (r=0.31, p=0.01), and divergent with self-reported avoidant behaviours (H1d, AAQ, r= -0.25, p=0.02). Hypotheses H1a and H1d were supported. However, as the association with the attention to feelings dimension of the TMMS was not significant (r= -0.15, p=0.18), our

emotional intelligence hypothesis (H1b) was only partially supported. Similarly, correlations with the empathy (H1c, IRI) dimensions were in the expected directions but were not significant: perspective taking (r=0.18, p=0.09), empathic concern (r=0.10, p=0.37), fantasy (r=0.17, p=0.11) and emotional distress (r= -0.15, p=0.16). H1c was therefore not supported.

< Insert Table 5 about here >

At subscale level, OMM data were positively associated with corresponding self-reports for the acceptance (r=0.35, p=0.01) and attentiveness items (r=0.29, p=0.01), but not the awareness item (r=0.15, p=0.17). Some variability was evident in the correlations between OMM subscales and the nomological net constructs. Of the three subscales, awareness had the strongest association with mindfulness (r=0.36, p=0.01), perspective taking (r=0.24, p=0.03), attention to feelings (r= -0.26, p=0.01) and clarity of feelings (r=0.25, p=0.02). The acceptance subscale was not convergent with self-reported mindfulness (r=12, p=0.25), although had stronger correlations with mood repair (r=0.35, p=0.01) and avoidance (r= -0.29, p=0.01) than the awareness or attentiveness subscales. The attentiveness subscale data was not significantly associated with any of the tested constructs.

Responsiveness

Baseline data showed good internal consistency for self-report mindfulness (MAAS α=0.91), observed mindfulness (OMM α=0.88) and the OMM subscales (awareness α=0.83, acceptance α=0.83, attentiveness α=0.78). Levene's test of equal variance showed no significant difference by group for either MAAS or OMM data. Shapiro Wilk's test for normality suggested the post-intervention OMM scores were not normally distributed, however the Q-Q plots and Kolmogorov-Smirnov tests supported normality, so transformations were not performed (Osborne, 2002). Nonparametric tests of difference were conducted to test the influence of normality violations on findings.

Box plots illustrate the raw change in observed mindfulness was higher for the MT group than the control (Figure 4), but the confidence intervals show high variability in the data.

< Figure 4 Title: Distribution of individual change scores for observed mindfulness.

Figure 4 Legend: OMM: Observed Mindfulness Measure. Mindfulness group n=32, Control group n=28 >

Table 6 presents the raw mean data by group and timepoint and the results of our tests of responsiveness (pre-post MT change). As hypothesized (H1a), change in observed mindfulness (t=0.94, p=0.35, Cohen's *d*=0.24, 95%CI -0.27, 0.75) was in the same direction but lower in magnitude than self-reported mindfulness (t=1.65, p=0.10, Cohen's *d*=0.42, 95%CI -0.10, 0.93). The results from nonparametric tests of difference supported these findings. Across the OMM subscales observed mindful awareness achieved a significant increase (t=2.02, p=0.05, Cohen's *d*=0.52, 95%CI 0.01, 1.04), while changes in mindful acceptance (t=0.15, p=0.88, Cohen's *d*=0.04, 95%CI -0.47, 0.55) and attentiveness (t= =0.15, p=0.89, Cohen's *d*= -0.04, 95%CI -0.54, 0.47) were negligible.

<Insert Table 6 about here>

Follow-up data suggested a further development of mindful acceptance beyond the training period, with a significant between groups difference in change from baseline to six-months for awareness (β=1.25, SE=0.44, p=0.01, *d*=0.29) and for acceptance (β=1.13, SE=0.43, p=0.01, *d*=0.33). Between groups change in attentiveness trended in the expected direction but remained non-significant at six-months (β=1.41, SE=1.03, p=0.17).

DISCUSSION

This paper presents the initial development of a new quantitative instrument for collecting second-person observations in mindfulness research, the Observed Mindfulness Measure (OMM). The nine-item OMM has three distinct, correlated dimensions: mindful

awareness, acceptance and attentiveness. The OMM has good psychometric properties and the ability to provide an overall score of observed mindful behaviours.

Scale development: construct definition, item selection and scale structure

Observed mindfulness was defined as the noticeable tendency of another person to be attentive to and aware of current experience, with an attitude of curiosity, openness and acceptance. Of a potential 51 items, 29 were considered to have face validity and were interpretable. A three-factor structure with a common latent variable was identified in exploratory factor analyses. This structure was confirmed in multiple independent samples comprising work, family and friend-based dyads, and after item responsiveness modelling, where poorly performing items were identified and dropped. The resultant nine-item measure was determined to have three dimensions. The labels, attentiveness, awareness and acceptance, were semantically derived from the items within each dimension and accord with our definition of observed mindfulness as well as key aspects of individual mindfulness as defined by others (Bishop et al., 2004). The multidimensional structure of the OMM supports prior work that has found mindfulness to have several interrelated components and was expected, given the dimensionality of measures from which the original items were drawn. The moderate correspondence of items and factors with the over-arching latent variable indicates it is appropriate to report a total summed score for the nine-items of the OMM for cross-validation purposes. However, differences that may be evident at factor level are likely to be obscured through pooling all items, so reporting OMM results at factor level is recommended for more in-depth enquiry.

Reliability and validity

The nine-item OMM showed excellent test-retest reliability. Concurrent criterion validity of the OMM was also supported with moderate to strong self-other agreement with an

established self-report mindfulness measure (MAAS; Brown and Ryan, 2003). In the absence of an existing gold standard second-person report measure, the MAAS was chosen as the primary criterion due to its wide acceptance and use as a valid measure of individual mindfulness (Qu et al., 2015). However, as this instrument has been criticised for not including the acceptance and non-judging dimensions of mindfulness assessed by other instruments (e.g. Grossman 2011), some discrepancy between self and other ratings on corresponding measures was expected. Method variance is another likely contributor to differences between self and second-person (observer) data, through expectancy effects, social desirability, priming, mood, or differences in the information available to the two types of respondent (Atkins and Wood, 2002; Olino and Klein, 2015; Ostroff et al., 2004; Roberts et al., 2013; Goldberg et al., 2019; Podsakoff et al., 2003). Vazire (2010) studied self-other knowledge asymmetry and found self-ratings to be more accurate than other-reports for internal experiences that may be low in observability, but other-ratings can be more accurate for more evaluative qualities such as actions and behaviours. The moderate-to-strong score for inter-rater agreement between the observer-reported OMM and self-reported MAAS data (0.45) indicates mindfulness is indeed a quality that manifests in behaviours that are noticeable to others, and supports the criterion validity of the OMM (DeVellis, 2012). This self-other agreement score exceeded the negligible correlation ($r=0.06$, ns) found by Bartlett et al. (2016) and the agreement between self- and other-reported versions of the FFMQ (0.19 to 0.25) reported by May and Reinhardt (2018). Thus, the purposefully developed second-person report OMM may provide a robust score that better represents the extent to which a known other person behaves mindfully, than has been obtained by adapting pre-existing self-report scales for other-report.

Construct validity of the OMM was partially supported. Emotional intelligence, empathy and acceptance were selected as key nomological network constructs for further validity testing because they are factors that influence social relationships and prior work has shown these

outcomes to be directly influenced by mindfulness (Aydin Sünbül, 2019; Donald et al., 2019; Jacobs et al., 2016; Block-Lerner et al., 2007). All tests showed the associations between the OMM and these constructs were in the expected direction, and no correlation was higher than 0.5, indicating the latent variable detected by the OMM is an independent construct

As hypothesized, the OMM data were convergent with self-reported emotional intelligence, and divergent with avoidance and emotional distress (DeVellis, 2012; Revelle and French, 2013). However, while the relationship of the OMM total score with perspective taking and empathic concern was in the hypothesized direction, correlations were small and not significant. This was unexpected given prior work showing clear associations of mindfulness with empathy and prosocial acting (Donald et al., 2019; Block-Lerner et al., 2007; Van Doesum et al., 2013). It is feasible the OMM items tap into a different, more externally focused aspect of mindfulness than is congruent with the construct of empathy represented in the IRI instrument. Perspective taking and empathic concern were hypothesized to correlate with OMM data based on literature review (e.g. Donald et al., 2019) and because these skills are theoretically cultivated through processes (reperceiving and decentering) involved in mindfulness meditation practice (Supplementary Figure S.1). These empathy skills may therefore be more strongly related to mindfulness meditation practice per se, than to self-reported dispositional mindfulness or observer-reported mindful behaviours.

The result may also be an artefact of measurement. Block-Lerner et al. (2007) noted that different self-report mindfulness instruments may have different relationships with the IRI after finding the Cognitive and Affective Mindfulness Scale (CAMS-R; Feldman et al., 2007) was significantly correlated with the perspective taking and empathic concern dimensions, but the Mindful Attention and Awareness Scale (MAAS; Brown and Ryan, 2003) was not. An ongoing discussion regarding variable conceptualisations of mindfulness in self-report instruments can be found elsewhere (e.g. Qu et al., 2015; Grossman, 2018; Bishop et al., 2004). This project

adapted the broadly accepted, composite definition of self-reported mindfulness proposed by Bishop et al. (2004), and an inclusive approach was taken to item selection, with the initial potential item set drawn from five different self-report mindfulness measures. Our results showed a significant correlation between the OMM awareness subscale and the IRI perspective taking dimension, but not empathic concern. This may be because perspective taking is considered the ability to "adopt another's psychological point of view" and is thus an externally oriented skill requiring awareness of that person's behaviours and reactions (Block-Lerner et al., 2007, p. 506), whereas empathic concern is closely associated with sympathy, which is a construct not represented in the conceptualisation of observed mindful behaviours as operationally defined.

In addition to the significant relationship with the perspective-taking dimension of empathy, the OMM awareness subscale data had differential relationships with the two feelings subscales in the emotional intelligence measure. Variability in the factor-level associations within the nomological net lends weight to the importance of assessing mindfulness as a multidimensional construct and supports the OMM as an instrument that may help increase our understanding of how a person's mindfulness might influence their behaviours, and potentially lead to external, social outcomes such as observed in previous studies (Horner et al., 2014; Cohen-Katz et al., 2005; Singh et al., 2015).

The acceptance factor had a stronger relationship with mood repair and avoidant behaviours, while attentiveness was not clearly convergent or divergent with any of the tested constructs. In previous research, attentional control clearly correlates with stress reduction and has been shown to precede awareness and acceptance as the acquisition of mindfulness skills develops (Garland et al., 2017b; Creswell and Lindsay, 2014). The absence of network associations between OMM attentiveness and the IRI distress scores flags a potential problem with the attentiveness subscale.

The responsiveness of the OMM to change was not confirmed. While the MT effects on observed mindful behaviours were as hypothesized, (i.e. in the same direction but lower in magnitude than self-reported mindfulness), neither self-reported or observer-reported mindfulness reached significance following mindfulness training. These null results prevented conclusions being drawn regarding the OMM's responsiveness. At sub-scale level, tests of change by group suggest observers of MT participants noticed an increase only in participants' mindful awareness immediately post-intervention, and that at follow-up mindful acceptance had developed further. The temporality of change in awareness and acceptance subscales support suggestions that mindfulness skills are cultivated sequentially, with mindful awareness preceding the development of mindful acceptance (Garland et al., 2017b). As well as showing poor convergence in construct validity testing, the attentiveness sub-scale data appeared to be subject to a ceiling effect, which limited the ability for changes to be detected.

The development of the OMM was premised on evidence of second-person reports providing accurate, informative and cost-effective research insights in the fields of personality, social psychology and educational research (Vazire, 2006; Olino and Klein, 2015). The OMM does not aim to assess the internal experiences of the subject being observed, rather it focuses on the degree to which the subject noticeably acts or responds mindfully in social contexts. OMM data collected to date indicate moderate-to-strong agreement between self-reported and observer-reported mindfulness. This supports the intended primary purpose of the OMM, to provide data for triangulating, and therefore supplementing, self-reported findings in mindfulness research. Used in conjunction with self-report measures, the OMM offers an easy to administer, scalable and inexpensive instrument to support further enquiry into the links between mindfulness and socially oriented behaviours in community and organizational settings. This quantitative measure can also potentially be used to help advance the growing field of research investigating the social effects of mindfulness. For example, OMM data may

help examine the differential influence of mindfulness dimensions on client engagement or work-based relationships, or the extent to which mindfulness can ameliorate the impact of work stress on familial relationships. Researching mindfulness as a malleable personal quality that yields outcomes with internal *and* external orientations may also help to align scientific narratives of this construct with its Buddhist roots, where mindfulness is conceived as part of a wholistic path to wisdom, mental discipline and moral conduct, not limited to the alleviation of personal suffering.

Limitations and future directions

Items in the initial item set were selected from uni- and multi-dimensional instruments based on both Eastern and Western definitions of mindfulness (Bergomi et al., 2013) with the intention of covering the spectrum of mindfulness conceptualisations. The final included items were drawn from the MAAS, KIMS, PHLMS and FMI measures. The approach employed was chosen for pragmatic reasons, and has been previously applied for the development of popular mindfulness measures such as the FFMQ (Baer et al., 2006). However, qualities of mindfulness that may be unique to the second-person view may have been missed. An alternative approach would have been to develop new items based on a phenomenological analysis of second-person descriptions of mindful qualities.

While the item and scale psychometrics for the OMM are sound, the attentiveness items did not show clear convergence with the nomological net constructs and appear to provide little information about skills at higher points on the scale, which may have led to the ceiling effect in our responsiveness tests. It is feasible the negative orientation of items in this subscale is problematic, since the other two subscales only include positively oriented items. Negative items, when mixed with positive items, can introduce an artefactual structure, random error and semantic confusion, although they can also help keep the attention of respondents and

make the scale less susceptible to social desirability (Lietz, 2010; Roszkowski and Soven, 2010). Further, for a second-person measure, observers may be more reluctant to score their observed person poorly on a negatively worded item (e.g. “the person doesn’t pay attention”) than on a positive item (e.g. “the person pays attention”). A subsequent study is therefore planned to assess the psychometric properties of the OMM using positive wording of the attentiveness items. This work may also strengthen the construct validity of the attentiveness dimension.

Tests of concurrent validity provided support for the OMM based on procedural and outcome variables in the mindfulness nomological network, however the range of relationships tested was not exhaustive. The MAAS is only one of a range of mindfulness measures, and future work should test the correspondence of the OMM with multi-dimensional self-report instruments. Testing the relationship between the OMM and internal, resource-related outcomes such as hope, optimism, self-efficacy and resilience (psychological capital), and externally oriented variables such as novelty seeking, ethical and citizenship behaviours will help clarify the relationship between mindfulness and socially oriented outcomes.

The influence of potential sources of bias (e.g. mood, desirability, expectancy) on OMM findings should also be examined, and researchers are encouraged to test the interaction of observers’ mindfulness, demographic characteristics, the frequency of dyadic interactions, cultural factors and the longevity and nature of respondents’ relationships (e.g. friend, family, work) as potential effect moderators. These data would also help inform patterns of missingness and support imputation for incomplete cases in future studies.

Finally, the RCT sample used for responsiveness testing enabled the comparison of training effects with an inactive control as well as tests of associations with other related constructs, which was expected to strengthen this stage of the study. However, the absence of significant changes in self-reported mindfulness following training limited assessable impacts

for observed mindfulness. Using the OMM in future studies where more robust changes in individual mindfulness are evidenced will help further understand the responsiveness and predictive utility of the OMM.

References

Allen, T. D., Eby, L. T., Conley, K. M., Williamson, R. L., Mancini, V. S., & Mitchell, M. E. (2015). What Do We Really Know About the Effects of Mindfulness-Based Training in the Workplace? *Industrial and Organizational Psychology*, 8(04), 652-661. doi:10.1017/iop.2015.95.

Atkins, P. W. B., & Wood, R. E. (2002). Self- Versus Others' Ratings As Predictors Of Assessment Center Ratings: Validation Evidence For 360-Degree Feedback Programs. *Personnel Psychology*, 55(4), 871-904. doi:10.1111/j.1744-6570.2002.tb00133.x.

Aydin Sünbül, Z. (2019). Indirect relations of mindful awareness with different forms of prosocial actions within college students: The mediating role of empathy. *Current Psychology*. doi:10.1007/s12144-019-00342-6.

Baer, R. A., Smith, G. T., & Allen, K. B. (2004). Assessment of Mindfulness by Self-Report: The Kentucky Inventory of Mindfulness Skills. *Assessment*, 11(3), 191-206. doi:10.1177/1073191104268029.

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(1), 27-45. doi:10.1177/1073191105283504.

Baer, R. A., Walsh, E., & Lykins, E. L. (2009). Assessment of Mindfulness. *Clinical Handbook of Mindfulness* (pp. 153-168). NY, USA: Springer.

Barnes, S., Brown, K. W., Krusemark, E., Campbell, W. K., & Rogge, R. D. (2007). The role of mindfulness in romantic relationship satisfaction and responses to relationship stress. *Journal of Marital and Family Therapy*, 33(4), 482-500. doi:10.1111/j.1752-0606.2007.00033.x.

Bartlett, L., Lovell, P., Otahal, P., & Sanderson, K. (2016). Acceptability, Feasibility, and Efficacy of a Workplace Mindfulness Program for Public Sector Employees: a Pilot Randomized Controlled Trial with Informant Reports. *Mindfulness*, 8(3), 639-654. doi:10.1007/s12671-016-0643-4.

Bartlett, L., Martin, A., Neil, A. L., Memish, K., Otahal, P., Kilpatrick, M., et al. (2019). A Systematic Review and Meta-Analysis of Workplace Mindfulness Training Randomized Controlled Trials. *Journal of Occupational Health Psychology*, 24(1), 108-126. doi:10.1037/ocp0000146.

Bartlett, L., Martin, A. J., Kilpatrick, M., Otahal, P., Sanderson, K., & Neil, A. L. (2021). Effects of a mindfulness app on employee stress: results of a randomised controlled trial in an Australian public sector workforce. *JMIR*, Pre-print. doi:10.2196/preprints.30272

Baumeister, R. F., Vohs, K. D., & Funder, D. C. (2007). Psychology as the Science of Self-Reports and Finger Movements: Whatever Happened to Actual Behavior? *Perspectives on Psychological Science*, 2(4), 396-403. doi:10.1111/j.1745-6916.2007.00051.x.

Beaujean, A. A. (2014). Latent Variable Modeling Using R. New York and London: Routledge Taylor & Francis Group.

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

Bernaards, C. A., & Jennrich, R. I. (2005). Gradient Projection Algorithms and Software for Arbitrary Rotation Criteria in Factor Analysis. *Educational and Psychological Measurement*, 65, 676-696.

Bishop, S. R., Lau, M., Shapiro, S., Carlson, L. E., Anderson, N. D., Carmody, J., et al. (2004). Mindfulness: A Proposed Operational Definition. *Clinical Psychology: Science and Practice*, 11(3), 230-241. doi:10.1093/clipsy/bph077.

Block-Lerner, J., Adair, C., Plumb, J. C., Rhatigan, D. L., & Orsillo, S. M. (2007). The case for mindfulness-based approaches in the cultivation of empathy: Does nonjudgmental, present-moment awareness increase capacity for perspective-taking and empathic concern? *Journal of Marital and Family Therapy*, 33(4), 501-516. doi:10.1111/j.1752-0606.2007.00034.x.

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(4), 822-848. doi:10.1037/0022-3514.84.4.822.

Cairns, V., & Murray, C. (2015). How Do the Features of Mindfulness-Based Cognitive Therapy Contribute to Positive Therapeutic Change? A Meta-Synthesis of Qualitative Studies. *Behavioural and Cognitive Psychotherapy*, 43(3), 342-359. doi:10.1017/S1352465813000945.

Cardaciotto, L., Herbert, J. D., Forman, E. M., Moitra, E., & Farrow, V. (2008). The Assessment of Present-Moment Awareness and Acceptance: The Philadelphia Mindfulness Scale. *Assessment*, 15(2), 204-223. doi:10.1177/1073191107311467.

Chalmers, R. P. (2012). mirt: A Multidimensional Item Response Theory Package for the R Environment. *2012*, 48(6), 29. doi:10.18637/jss.v048.i06.

Chambers, R., Gullone, E., & Allen, N. B. (2009). Mindful emotion regulation: An integrative review. *Clinical Psychology Review*, 29(6), 560-572. doi:10.1016/j.cpr.2009.06.005.

Chiesa, A., Calati, R., & Serretti, A. (2011). Does mindfulness training improve cognitive abilities? A systematic review of neuropsychological findings. *Clinical Psychology Review*, 31(3), 449-64. doi:10.1016/j.cpr.2010.11.003.

Chiesa, A., & Serretti, A. (2010). A systematic review of neurobiological and clinical features of mindfulness meditations. *Psychological Medicine*, 40(8), 1239-52. doi:10.1017/s0033291709991747.

Choi, E., & Leroy, H. (2015). Methods of mindfulness: How mindfulness is studied in the workplace. *Mindfulness in organizations: Foundations, research, and applications.* (pp. 67-99). Cambridge University Press, New York, NY. doi:10.1017/9781107587793.006.

Cohen-Katz, J., Wiley, S., Capuano, T., Baker, D. M., Deitrick, L., & Shapiro, S. (2005). The effects of mindfulness-based stress reduction on nurse stress and burnout: a qualitative and quantitative study, part III. *Holistic Nursing Practice*, 19(2), 78-86.

Connelly, B. S., & Ones, D. S. (2010). An Other Perspective on Personality: Meta-Analytic Integration of Observers' Accuracy and Predictive Validity. *Psychological Bulletin*, 136(6), 1092-1122.

Creswell, J. D., & Lindsay, E. K. (2014). How Does Mindfulness Training Affect Health? A Mindfulness Stress Buffering Account. *Current Directions in Psychological Science*, 23(6), 401-407. doi:10.1177/0963721414547415.

Cronbach, L., & Meehl, P. (1955). Construct validity in psychological tests. *Psychological Bulletin*, 52(4), 281-302.

Davidson, R. J., & Kaszniak, A. W. (2015). Conceptual and methodological issues in research on mindfulness and meditation. *American Psychologist*, 70(7), 581-592. doi:10.1037/a0039512.

Davis, M. H. (1983). Measuring individual differences in empathy: Evidence for a multidimensional approach. *Journal of Personality and Social Psychology*, 44(1), 113.

DeVellis, R. F. (2012). Scale Development: Theory and Applications (Applied Social Research Methods). London UK: Sage Publications.

DiStefano, C., Zhu, M., & Mindrila, D. (2009). Understanding and Using Factor Scores: Considerations for the Applied Researcher. *Practical Assessment, Research & Evaluation*, 14(20), 1-11.

Donald, J. N., Sahdra, B. K., Van Zanden, B., Johannes Duineveld, J., Atkins, P. W. B., Marshall, S., et al. (2019). Does Your Mindfulness Benefit Others? A Systematic Review and Meta-Analysis of the Link Between Mindfulness and Prosocial Behavior. *British Journal of Psychology*, 110(1), 101-125. doi:10.1111/bjop.12338.

Dong, Y., & Peng, C.-Y. J. (2013). Principled missing data methods for researchers. *SpringerPlus*, 2(1), 222-222. doi:10.1186/2193-1801-2-222.

Duncan, L. G. (2007) Assessment of Mindful Parenting Among Parents of Early Adolescents: Development and Validation of the Interpersonal Mindfulness in Parenting Scale. PhD, Pennsylvania State University

Dunn, T. J., Baguley, T., & Brunsden, V. (2014). From alpha to omega: A practical solution to the pervasive problem of internal consistency estimation. *British Journal of Psychology*, 105(3), 399-412. doi:10.1111/bjop.12046.

Eby, L. T., Allen, T. D., Conley, K. M., Williamson, R. L., Henderson, T. G., & Mancini, V. S. (2017). Mindfulness-based training interventions for employees: A qualitative review of the literature. *Human Resource Management Review*. doi:10.1016/j.hrmr.2017.03.004.

Eisenberg, N. (2015). Prosocial Emotion: Methods and Current Research. In N. Eisenberg (Ed.), *Altruistic Emotion, Cognition and Behavior* (pp. 30-57). London, UK: Taylor & Francis Group. http://ebookcentral.proquest.com/lib/utas/detail.action?docID=1864870.

Feldman, G., Hayes, A., Kumar, S., Greeson, J., & Laurenceau, J.-P. (2007). Mindfulness and emotion regulation: The development and initial validation of the Cognitive and Affective Mindfulness Scale-Revised (CAMS-R). *Journal of Psychopathology and Behavioral Assessment*, 29(3), 177-190.

Floyd, R. G. (2005). Measurement Properties of Indirect Assessment Methods for Functional Behavioral Assessment: A Review of Research. *School Psychology Review*, 34(1), 58-73.

Fu, R., Vandermeer, B. W., Shamliyan, T. A., O'Neil, M. E., Yazdi, R., Fox, S. H., et al. (2013). Handling Continuous Outcomes in Quantitative Synthesis. *Methods Guide for Effectiveness and Comparative Effectiveness Reviews* Rockville, MD: Agency for Healthcare Research and Quality. www.ncbi.nlm.nih.gov/books/NBK154408/.

Funder, D. C., & Dobroth, K. M. (1987). Differences Between Traits: Properties Associated with Interjudge Agreement. *Journal of Personality & Social Psychology*, 52(2), 409-418.

Funder, D. C., & West, S. G. (1993). Consensus, Self-Other Agreement, and Accuracy in Personality Judgment: An Introduction. *Journal of Personality*, 61(4), 457-476. doi:10.1111/j.1467-6494.1993.tb00778.x.

Garland, E. L., Hanley, A. W., Baker, A. K., & Howard, M. O. (2017a). Biobehavioral Mechanisms of Mindfulness as a Treatment for Chronic Stress: An RDoC Perspective. *Chronic Stress*, 1, 1–14. doi:10.1177/2470547017711912.

Garland, E. L., Hanley, A. W., Goldin, P. R., & Gross, J. J. (2017b). Testing the mindfulness-to-meaning theory: Evidence for mindful positive emotion regulation from a ranalysis of longitudinal data. *PLoS ONE*, 12(12), e0187727. doi:10.1371/journal.pone.0187727.

Gillespie, B., Davey, M. P., & Flemke, K. (2015). Intimate Partners' Perspectives on the Relational Effects of Mindfulness-Based Stress Reduction Training: A Qualitative Research Study. *Contemporary Family Therapy*, 37(4), 396-407. doi:10.1007/s10591-015-9350-x.

Goldberg, S. B., Tucker, R. P., Greene, P. A., Simpson, T. L., Hoyt, W. T., Kearney, D. J., et al. (2019). What Can We Learn from Randomized Clinical Trials About the Construct Validity of Self-Report Measures of Mindfulness? A Meta-Analysis. *Mindfulness*, 10(5), 775-785. doi:10.1007/s12671-018-1032-y.

Goldberg, S. B., Tucker, R. P., Greene, P. A., Simpson, T. L., Kearney, D. J., & Davidson, R. J. (2017). Is mindfulness research methodology improving over time? A systematic review. *PLOS ONE*, 12(10). doi:10.1371/journal.pone.0187298.

Good, D. J., Lyddy, C. J., Glomb, T. M., Bono, J. E., Brown, K. W., Duffy, M. K., et al. (2016). Contemplating Mindfulness at Work. *Journal of Management*, 42(1), 114. doi:10.1177/0149206315617003.

Grossman, P. (2018). On the porosity of subject and object in "mindfulness" scientific study: Challenges to "scientific" construction, operationalization and measurement of mindfulness. *MindRxiv*. doi:10.31231/osf.io/3hqg8.

Grossman, P., & Van Dam, N. T. (2011). Mindfulness, by any other name…: trials and tribulations of sati in western psychology and science. *Contemporary Buddhism*, 12(1), 219-239. doi:10.1080/14639947.2011.564841.

Harris, P. A., Taylor, R., Thielke, R., Payne, J., NGonzalez, N., & Conde, J. G. (2009). Research electronic data capture (REDCap) - A metadata-driven methodology and workflow process for providing translational research informatics support. *Journal of Biomedical Information*, 42(2), 377-381.

Hayes, S. C., Luoma, J. B., Bond, F. W., Masuda, A., & Lillis, J. (2006). Acceptance and commitment therapy: Model, processes and outcomes. *Behaviour Research and Therapy*, 44(1), 1-25.

Hays, R. D., Morales, L. S., & Reise, S. P. (2000). Item response theory and health outcomes measurement in the 21st century. *Medical Care*, 38(9), II28-II42.

Horner, J. K., Piercy, B. S., Eure, L., & Woodard, E. K. (2014). A pilot study to evaluate mindfulness as a strategy to improve inpatient nurse and patient experiences. *Applied Nursing Research*, 27(3), 198.

Hugh-Jones, S., Rose, S., Koutsopoulou, G. Z., & Simms-Ellis, R. (2018). How Is Stress Reduced by a Workplace Mindfulness Intervention? A Qualitative Study Conceptualising Experiences of Change. *Mindfulness*, 9(2), 474-487. doi:10.1007/s12671-017-0790-2.

Hyland, P. K., Lee, R. A., & Mills, M. J. (2015). Mindfulness at Work: A New Approach to Improving Individual and Organizational Performance. *Industrial and Organizational Psychology*, 8(04), 576-602. doi:10.1017/iop.2015.41.

Jacobs, I., Wollny, A., Sim, C. W., & Horsch, A. (2016). Mindfulness facets, trait emotional intelligence, emotional distress, and multiple health behaviors: A serial two-mediator model. *Scandinavian Journal of Psychology*, 57(3), 207-14. doi:10.1111/sjop.12285.

Jaurequi, M. E. (2019). Adult Attachment and Mindfulness on Mental Health: A Systematic Research Synthesis. *Journal of Relationships Research*, 10, e17. doi:10.1017/jrr.2019.15.

Jha, A. P., Stanley, E. A., Kiyonaga, A., Wong, L., & Gelfand, L. (2010). Examining the protective effects of mindfulness training on working memory capacity and affective experience. *Emotion*, 10(1), 54-64. doi:10.1037/a0018438.

Karremans, J. C., Schellekens, M. P., & Kappen, G. (2017). Bridging the Sciences of Mindfulness and Romantic Relationships. *Personality and Social Psychology Review*, 21(1), 29-49. doi:10.1177/1088868315615450.

Khaddouma, A., Coop Gordon, K., & Strand, E. B. (2017). Mindful Mates: A Pilot Study of the Relational Effects of Mindfulness-Based Stress Reduction on Participants and Their Partners. *Family Process*, 56(3), 636-651. doi:10.1111/famp.12226.

Khoury, B., Sharma, M., Rush, S. E., & Fournier, C. (2015). Mindfulness-based stress reduction for healthy individuals: A meta-analysis. *Journal of Psychosomatic Research*, 78(6), 519-528. doi:10.1016/j.jpsychores.2015.03.009.

Kirkwood, B. R., & Sterne, J. A. C. (2013). Essential Medical Statistics. Oxford, UK: Blackwell.

Koo, T. K., & Li, M. Y. (2016). A Guideline of Selecting and Reporting Intraclass Correlation Coefficients for Reliability Research. *Journal of Chiropractic Medicine*, 15(2), 155-163. doi:10.1016/j.jcm.2016.02.012.

Kreplin, U., Farias, M., & Brazil, I. A. (2018). The limited prosocial effects of meditation: A systematic review and meta-analysis. *Scientific Reports*, 8(1), 2403. doi:10.1038/s41598-018-20299-z.

Lemberger-Truelove, M. E., Carbonneau, K. J., Zieher, A. K., & Atencio, D. J. (2019). Support for the Development and Use of the Child Observation of Mindfulness Measure (C-OMM). *Mindfulness*. doi:10.1007/s12671-019-1094-5.

Lewin, S., Glenton, C., Munthe-Kaas, H., Carlsen, B., Colvin, C. J., Gülmezoglu, M., et al. (2015). Using Qualitative Evidence in Decision Making for Health and Social Interventions: An Approach to Assess Confidence in Findings from Qualitative Evidence Syntheses (GRADE-CERQual). *PLOS Medicine*, 12(10), e1001895. doi: 10.1371/journal.pmed.1001895.

Lietz, P. (2010). Research into questionnaire design: A summary of the literature. *International Journal of Market Research*, 52(2). doi:10.2501/S147078530920120X.

Madill, A., & Gough, B. (2008). Qualitative Research and its Place in Psychological Science. *Psychological Methods*, 13(3), 254-271.

Malinowski, P. (2013). Neural mechanisms of attentional control in mindfulness meditation (Focused Review). *Frontiers in Neuroscience*, 7(8). doi:10.3389/fnins.2013.00008.

Markon, K. E., Quilty, L. C., Bagby, R. M., & Krueger, R. F. (2013). The development and psychometric properties of an informant-report form of the personality inventory for DSM-5 (PID-5). *Assessment*, 20(3), 370-383. doi:10.1177/1073191113486513.

May, L. M., & Reinhardt, K. M. (2018). Self-Other Agreement in the Assessment of Mindfulness Using the Five-Facet Mindfulness Questionnaire. *Mindfulness*, 9(1), 105-116. doi:10.1007/s12671-017-0749-3.

Miao, C., Humphrey, R. H., & Qian, S. (2018). The relationship between emotional intelligence and trait mindfulness: A meta-analytic review. *Personality and Individual Differences*, 135, 101-107. doi:10.1016/j.paid.2018.06.051.

Muraki, E., & Muraki, M. (2016). Generalized Partial Credit Model. In W. J. van der Linden (Ed.), *Handbook of Item Response Theory* New York: Taylor & Francis.

Olino, T. M., & Klein, D. N. (2015). Psychometric Comparison of Self- and Informant-Reports of Personality. *Assessment*, 22(6), 655-664. doi:10.1177/1073191114567942.

Osborne, J. W. (2002). Notes on the use of data transformations. *Practical Assessment, Research & Evaluation*, 8. doi:10.7275/4vng-5608

Ostroff, C., Atwater, L. E., & Feinberg, B. J. (2004). Understanding Self-Other Agreement: A Look at Rater and Ratee Characteristics, Context and Outcomes. *Personnel Psychology*, 57(2), 333-375. doi:doi:10.1111/j.1744-6570.2004.tb02494.x.

Pirson, M., Langer, E., & Zilcha, S. (2018). Enabling a Socio-cognitive Perspective of Mindfulness: The Development and Validation of the Langer Mindfulness Scale. *Journal of Adult Development*, 25, 168-185. doi:10.1007/s10804-018-9282-4.

Podsakoff, P. M., MacKenzie, S. B., Lee, J.-Y., & Podsakoff, N. P. (2003). Common Method Biases in Behavioral Research: A Critical Review of the Literature and Recommended Remedies. *Journal of Applied Psychology*, 88(5), 879-903. doi:10.1016/j.paid.2018.06.051.

Pratscher, S. D., Wood, P. K., King, L. A., & Bettencourt, B. A. (2018). Interpersonal Mindfulness: Scale Development and Initial Construct Validation. *Mindfulness*. doi:10.1007/s12671-018-1057-2.

Qu, Y., Dasborough, M. T., & Todorova, G. (2015). Which Mindfulness Measures To Choose To Use? *Industrial and Organizational Psychology*, 8(04), 710-723. doi:10.1017/iop.2015.105.

R Core Team (2014) 'R: A language and environment for statistical computing.'. R Foundation for Statistical Computing. Vienna, Austria. Available at: www.R-project.org.

Reb, J., & Choi, E. (2014). Mindfulness in organizations. *Psychology of Meditation.* (pp. 279-309). Nova Science Publishers, Hauppauge, NY.

Revelle, W. (2019). psych: Procedures for Psychological, Psychometric, and Personality Research. *The Comprehensive R Archive Network (CRAN)*, Northwestern University, Evanston, Illinois. www.CRAN.R-project.org/package=psych.

Revelle, W., & French, J. A. (2013). The Personality Project. Personality, Motivation and Cognition Laboratory, Northwestern University. https://www.personality-project.org. Accessed 2018.

Rizopoulos, D. (2006). ltm: An R package for Latent Variable Modelling and Item Response Theory Analyses. *Journal of Statistical Software*, 17(5), 1-25.

Roberts, M. J., Campbell, J. L., Richards, S. H., & Wright, C. (2013). Self-Other Agreement in Multisource Feedback: The Influence of Doctor and Rater Group Characteristics. *Journal of Continuing Education in the Health Professions*, 33(1), 14-23. doi:10.1002/chp.21162.

Roche, M., Haar, J. M., & Luthans, F. (2014). The role of mindfulness and psychological capital on the well-being of leaders. *Journal of Occupational Health Psychology*, 19(4), 476. doi:10.1037/a0037183.

Rosenthal, R. (2009). Experimenter Effects in Behavioral Research. In R. Rosenthal, & R. L. Rosnow (Eds.), *Artifacts in Behavorial Research* (pp. 297-652). NY, USA: Oxford University Press.

Rosseel, Y. (2012). lavaan: An R Package for Structural Equation Modeling. *Journal of Statistical Software*, 48(2), 1-36.

Roszkowski, M. J., & Soven, M. (2010). Shifting gears: consequences of including two negatively worded items in the middle of a positively worded questionnaire. *Assessment & Evaluation in Higher Education*, 35(1), 117-134. doi:10.1080/02602930802618344.

Salovey, P., Mayer, J. D., Goldman, S. L., Turvey, C., & Palfai, T. P. (1995). Emotional attention, clarity, and repair: Exploring emotional intelligence using the Trait Meta-Mood Scale. In J. W. Pennebaker (Ed.), *Emotion, Disclosure, and Health* (pp. 125-154). Washington DC: American Psychological Association.

Sandelowski, M., Docherty, S., & Emden, C. (1997). Qualitative Metasynthesis: Issues and Techniques. *Research in Nursing & Health*, 20(4), 365-371. doi:10.1002/.

Sharp, J. E., & Jennings, P. A. (2016). Strengthening teacher presence through mindfulness: What educators say about the Cultivating Awareness and Resilience in Education (CARE) program. *Mindfulness*, 7(1), 209-218. doi:10.1007/s12671-015-0474-8.

Singh, N. N., Lancioni, G. E., Karazsia, B. T., Myers, R. E., Winton, A. S. W., Latham, L. L., et al. (2015). Effects of training staff in MBPBS on the use of physical restraints, staff stress and turnover, staff and peer injuries, and cost effectiveness in developmental disabilities. *Mindfulness*, 6(4), 926-937. doi:10.1007/s12671-011-0056-3.

Singh, N. N., Lancioni, G. E., Winton, A. S. W., Singh, J., Singh, A. N., Adkins, A. D., et al. (2010). Training in Mindful Caregiving Transfers to Parent–Child Interactions. *Journal of Child and Family Studies*, 19(2), 167-174. doi:10.1007/s10826-009-9267-9.

Siu, A. F. Y., Ma, Y., & Chui, F. W. Y. (2016). Maternal Mindfulness and Child Social Behavior: the Mediating Role of the Mother-Child Relationship. *Mindfulness*, 7(3), 577-583. doi:10.1007/s12671-016-0491-2.

Smith, G. T., Fischer, S., & Fister, S. M. (2003). Incremental Validity Principles in Test Construction. *Psychological Assessment*, 15(4), 467-477. doi:10.1037/1040-3590.15.4.467.

Taylor, C., Harrison, J., Haimovitz, K., Oberle, E., Thomson, K., Schonert-Reichl, K., et al. (2016). Examining ways that a mindfulness-based intervention reduces stress in public school teachers: A mixed-methods study. *Mindfulness*, 7(1), 115-129. doi:10.1007/s12671-015-0425-4.

Thera, S. (1998). The Way of Mindfulness: The Satipatthana Sutta and Its Commentary. Access to Insight. https://www.accesstoinsight.org/lib/authors/soma/wayof.html. Accessed 30 January 2019.

Van Dam, N. T., van Vugt, M. K., Vago, D. R., Schmalzl, L., Saron, C. D., Olendzki, A., et al. (2017). Mind the Hype: A Critical Evaluation and Prescriptive Agenda for Research on Mindfulness and Meditation. *Perspectives on Psychological Science*, 13(1), 36-61. doi:10.1177/1745691617709589.

Van de Mortel, T. F. (2008). Faking it: social desirability response bias in self-report research. *The Australian Journal of Advanced Nursing*, 25(4), 40-48.

Van Doesum, N. J., Karremans, J. C., Fikke, R. C., de Lange, M. A., & Van Lange, P. A. M. (2018). Social mindfulness in the real world: the physical presence of others induces other-regarding motivation. *Social Influence*, 13(4), 209-222. doi:10.1080/15534510.2018.1544589.

Van Doesum, N. J., Van Lange, D. A. W., & Van Lange, P. A. M. (2013). Social mindfulness: Skill and will to navigate the social world. *Journal of Personality and Social Psychology*, 105(1), 86-103. doi:10.1037/a0032540.

Van Lange, P. A., & Van Doesum, N. J. (2015). Social mindfulness and social hostility. *Current Opinion in Behavioral Sciences*, 3, 18-24. doi:10.1016/j.cobeha.2014.12.009.

Varela, F. J., & Shear, J. (1999). First-person Methodologies: What, Why, How? *Journal of Consciousness Studies*, 6(2-3), 1-14.

Vazire, S. (2006). Informant reports: A cheap, fast, and easy method for personality assessment. *Journal of Research in Personality*, 40(5), 472-481. doi:10.1016/j.jrp.2005.03.003.

Vazire, S. (2010). Who Knows What About a Person? The Self-Other Knowledge Asymmetry (SOKA) Model. *Journal of Personality & Social Psychology*, 98(2), 281-300. doi:10.1037/a0017908.

Vazire, S., & Mehl, M. R. (2008). Knowing Me, Knowing You: The Accuracy and Unique Predictive Validity of Self-Ratings and Other-Ratings of Daily Behavior. *Journal of Personality & Social Psychology*, 95(5), 1202-1216. doi:10.1037/a0013314.

Virgili, M. (2015). Mindfulness-Based Interventions Reduce Psychological Distress in Working Adults: a Meta-Analysis of Intervention Studies. *Mindfulness*, 6(2), 326-337. doi:10.1007/s12671-013-0264-0.

Walach, H., Buchheld, N., Buttenmüller, V., Kleinknecht, N., & Schmidt, S. (2006). Measuring mindfulness—the Freiburg Mindfulness Inventory (FMI). *Personality and Individual Differences*, 40(8), 1543-1555. doi:10.1016/j.paid.2005.11.025.

Table 1. Participant characteristics for Samples A, B and C

Sample A			Sample B				Sample C				
	Observers			Participants		Observers		Participants		Observers	
Complete cases	n=307		Complete cases	n=88		n=88	Complete cases	n=99		n=99	
	n	%		n	%	n	%	n	%	n	%
Age (yrs)			Age (yrs)				Age (yrs)				
18 to 24	10	(3)	18 to 24	4	(5)	3	(3)	0	0	0	0
25 to 34	30	(10)	25 to 34	11	(13)	14	(16)	8	(8)	15	(15)
35 to 44	68	(22)	35 to 44	21	(24)	19	(22)	28	(28)	31	(31)
45 to 54	109	(36)	45 to 54	26	(30)	23	(26)	36	(36)	29	(29)
55 to 64	66	(22)	55 to 64	15	(17)	18	(21)	26	(26)	21	(21)
Over 65	24	(8)	Over 65	11	(13)	11	(13)	1	(1)	3	(3)
Female	233	(76)	Female	54	(61)	49	(56)	76	(77)	72	(73)
Observer’s relationship			Observer's relationship					Observer's relationship			
Colleague	36	(9)	Colleague			10	(11)	Co-worker		43	(44)
Friend	75	(18)	Friend			24	(27)	Work friend		17	(17)
Family	150	(37)	Family			54	(61)	Subordinate		21	(21)
								Supervisor		12	(12)
								Other		5	(5)

Table 2. Step-wise item reduction for the 30 items endorsed by Expert Advisors for potential inclusion in the Observed Mindfulness Measure

Item number	Items and instructions (as presented to Sample A)	% "0" responses	Retained for EFA	Retained following EFA	Flagged for exclusion following IRT	Included in final measure
	Thinking of somebody you know well, please tell us how frequently each of the following statements would be true, where 1 = Not at all, 2 = Occasionally, 3 = About half the time, 4 = Frequently, and 5 = All the time. If none of these response options is appropriate, please select 0 = Cannot answer and provide a brief reason in the popup box provided.	Stage 2			Stage 3	
OMM1	The person has many mishaps or accidents	0.65	✓			
OMM2	The person has difficulty staying focused on what is happening to/around them as it occurs	0.33	✓	✓		✓
OMM3	The person seems to 'run on automatic' without much awareness of what he or she is doing	0.98	✓	✓		✓
OMM4	The person does other things while listening to someone	0.00	✓	✓	✓	
OMM5	The person seems preoccupied with the future or the past	0.33	✓			
OMM6	The person is good at finding the words to describe his or her feelings	1.30	✓	✓	✓	
OMM7	The person can express his or her beliefs, opinions and expectations	0.00	✓	✓	✓	
OMM8	The person can describe how things taste, smell or sound	0.65	✓	✓	✓	
OMM9	The person doesn't pay attention to what he or she is doing, because of daydreaming, worrying or other distractions	0.65	✓	✓		✓
OMM10	The person makes judgments about how worthy his or her experiences are	3.58				
OMM11	Even when terribly upset, this person can find words to explain what is going on	1.30	✓	✓		
OMM12	The person seems aware of how emotions affect his or her thoughts and behaviour	0.65	✓	✓	✓	✓
OMM13	When asked how he or she is feeling, the person can identify their emotions easily	0.98	✓	✓		✓
OMM14	The person seems aware of his or her own emotions when interacting with others	0.00	✓	✓	✓	✓
OMM15	The person notices what other people are up to	0.65	✓		✓	
OMM16	The person avoids thought provoking conversations	0.33	✓			
OMM17	The person tries to think of new ways of doing things	0.33	✓			
OMM18	The person is alert to new developments	0.65	✓			
OMM19	The person appears to appreciate him- or her self	0.33	✓	✓	✓	
OMM20	The person appears to accept his or her mistakes and difficulties without judging them	0.33	✓	✓	✓	
OMM21	The person seems to recover well from unpleasant or stressful experiences	0.33	✓	✓	✓	✓

Table 2. Step-wise item reduction for the 30 items endorsed by Expert Advisors for potential inclusion in the Observed Mindfulness Measure

Item number	Items and instructions (as presented to Sample A)	% "0" responses	Retained for EFA	Retained following EFA	Flagged for exclusion following IRT	Included in final measure
OMM22	The person gets upset with him- or herself when things go wrong	0.65	✓			
OMM23	The person can pause before reacting to difficult situations	0.65	✓	✓	✓	✓
OMM24	The person remains calm, even when things get hectic and stressful	0.00	✓	✓	✓	✓
OMM25	The person is impatient with him or herself and with others	0.00	✓			
OMM26	The person pays attention to what I am saying when we are talking	0.00	✓			
OMM27	The person welcomes new ideas	0.33	✓			
OMM28	The person becomes upset when faced with changing circumstances	0.33	✓			
OMM29	The person avoids difficult situations	0.98	✓			
OMM30	The person holds onto past experiences	1.63	✓	✓	✓	

EFA: Exploratory Factor Analysis; IRT: Item Response Theory Modelling; OMM: Observed Mindfulness Measure
Retained for EFA if >2% respondents could not provide an answer to this item; Retained after EFA if item loads >0.5 on one factor with at least 0.3 difference in loadings on other factors; Flagged for exclusion following IRT if item showed poor discrimination, did not detect a range of difficulty/ability or was potentially redundant.

Table 3. Item responsiveness test results and item reduction decisions for the Observed Mindfulness Measure (18 items)

Factor label	Item	Keep	Reason for decision	Potential redundancy	Difficulty thresholds per response option				Item discrimination		Item-factor loading	Scoring direction
				χ^2 p-value	*b*1	*b*2	*b*3	*b*4	Trace curves	*a*-value	< 0.5	
Attentiveness	omm2	keep	good item, detects range of ability	0.444	-3.039	-1.333	-1.099	0.273	✓	1.602	x	Reverse
	omm3	keep	good item, detects range of ability	0.414	-2.177	-1.195	-0.788	0.290	✓	2.416	x	Reverse
	omm4	discard	poor discrimination	0.766	-2.448	-0.732	-1.121	1.575	x	0.551	✓	Reverse
	omm9	keep	reasonable item, detects low ability	0.357	-1.972	-1.138	-1.214	-0.144	✓	1.279	x	Reverse
Awareness	omm6	discard	potential duplication	0.314	-1.24	-0.639	-0.598	1.39	✓	1.211	x	
	omm7	discard	potential duplication	0.308	-2.551	-1.415	-1.237	0.656	✓	1.069	x	
	omm8	discard	poor discrimination, difficulty skewed	0.731	-2.426	-1.465	-1.276	0.553	x	0.849	✓	
	omm11	discard	item fit, difficulty skewed	0.024	-1.239	-0.561	-0.607	1.467	x	1.407	x	
	omm12	keep	good item, detects range of ability	0.688	-1.382	-0.685	-0.295	1.307	✓	1.826	x	
	omm13	keep	item fit, difficulty in low range	0.035	-1.691	-0.79	-0.282	0.993	✓	2.978	x	
	omm14	keep	potential duplication	0.542	-2.023	-0.789	-0.208	1.157	✓	2.107	x	
	omm15	discard	poor discrimination	0.279	-2.164	-1.104	-0.741	1.615	x	0.907	✓	
Acceptance	omm19	discard	poor discrimination	0.252	-2.191	-1.026	-0.502	2.266	x	0.825	✓	
	omm20	discard	potential duplication	0.399	-1.357	-0.516	0.123	1.813	✓	1.653	x	
	omm21	keep	good item, detects range of ability	0.000	-1.341	-0.567	-0.105	1.702	✓	1.774	x	
	omm23	keep	good item, detects range of ability	0.202	-1.45	-0.604	-0.047	1.824	✓	1.259	x	
	omm24	keep	good item, detects range of ability	0.102	-1.297	-0.521	0.128	1.716	✓	1.264	x	
	omm30	discard	poor discrimination	0.283	-1.427	-0.577	-0.485	1.880	x	0.654	✓	Reverse

Item responsiveness tests conducted at factor (unidimensional scale) level using generalised partial credit model for ordinal, polytomous data. χ^2 p-value: significance of difference between predicted and observed response distributions for each item (α<0.05). *b*1, *b*2, *b*3, *b*4: threshold parameters for the point on the scale of the latent trait with 0.5 positive response probability. *a*-value: discrimination parameter showing item's ability to discriminate between different levels of the underlying trait relative to the threshold parameter. Trace curves: ✓/x = satisfactory/unsatisfactory item characteristic curves (ICCs) depicting relationship between the item and the latent trait.

Table 4. Fit statistics for the nine-item Observed Mindfulness Measure (OMM)

	Observer sample	Sample size (n)	No. factors	Model fit statistics				
				RMSEA	90%CI	CFI	TLI	χ^2 p
Confirmatory analyses	B+C	187	1	0.122	0.112, 0.133	0.723	0.686	<0.01
	B+C	187	3	0.019	0.000, 0.060	0.998	0.997	0.364
Sensitivity analyses	A	307	3	0.044	0.013, 0.069	0.987	0.980	0.033
	B	88	3	0.047	0.000, 0.103	0.984	0.976	0.230
	C	99	3	0.053	0.000, 0.096	0.984	0.977	0.111
	A+B+C	494	3	0.042	0.023, 0.060	0.992	0.987	0.005

Fit indices from maximum likelihood confirmatory factor analyses with one and three factors specified; RMSEA: Root Mean Square Error of Approximation & 95% Confidence Intervals; CFI: Comparative Fit Index; TLI: Tucker Lewis Index; Samples A and B: mix of work, friend and family observers; Sample C: work-based observers

Table 5. Associations between the Observed Mindfulness Measure (OMM), self-reported mindfulness, emotional intelligence, empathy and acceptance

	Observer-reported variables				Self-reported variables								
Construct variables	OMM total	OMM Awareness	OMM Acceptance	OMM Attentiveness	Mindfulness	Avoidance	Attention to feelings	Clarity of feelings	Mood repair	Perspective taking	Fantasy	Empathic concern	Psychological distress
OMM total		0.00	0.00	0.00	0.01	0.02	0.18	0.01	0.00	0.09	0.20	0.37	0.16
OMM Awareness	0.81*		0.00	0.00	0.00	0.07	0.01	0.02	0.17	0.03	0.11	0.08	0.74
OMM Acceptance	0.79*	0.42*		0.01	0.25	0.01	0.52	0.10	0.00	0.17	0.79	0.81	0.17
OMM Attentiveness	0.66*	0.41*	0.27*		0.07	0.68	0.08	0.06	0.11	0.94	0.21	0.48	0.09
Mindfulness (MAAS)	0.29*	0.36*	0.12	0.20		0.00	0.45	0.00	0.06	0.00	0.29	0.67	0.00
Avoidance (AAQ)	-0.25*	-0.19	-0.29*	-0.04	-0.39*		0.44	0.00	0.00	0.00	0.27	0.71	0.00
Attention to feelings (TMMS)	-0.15	-0.26*	0.07	-0.18	-0.08	0.08		0.01	0.83	0.26	0.00	0.00	0.64
Clarity of feelings (TMMS)	0.27*	0.25*	0.17	0.20	0.33*	-0.49*	-0.26*		0.01	0.83	0.26	0.00	0.00
Mood repair (TMMS)	0.31*	0.15	0.35*	0.17	0.20	-0.42*	-0.02	0.24*		0.03	0.74	0.20	0.01
Perspective taking (IRI)	0.18	0.24*	0.15	-0.01	0.30*	-0.33*	-0.12	0.17	0.23*		0.23	0.00	0.00
Fantasy (IRI)	0.14	0.17	0.03	0.14	-0.11	0.12	-0.41*	0.03	-0.03	0.13		0.00	0.33
Empathic concern (IRI)	0.10	0.18	-0.03	0.08	0.05	-0.04	-0.38*	0.14	0.13	0.38*	0.35*		0.98
Psychological distress (IRI)	-0.15	-0.04	-0.15	-0.18	-0.43*	0.57*	-0.05	-0.38*	-0.26*	-0.33*	0.10	0.00	

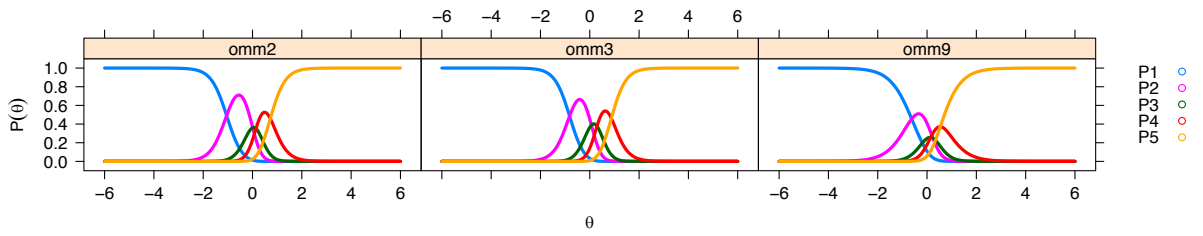
Shaded area: Pearson's product moment correlations. Unshaded area: corresponding p-values. * significant with α=0.05. Data: Sample B baseline n=88 dyads. OMM: Observed Mindfulness Measure; MAAS: Mindful Attention and Awareness Scale; AAQ: Acceptance and Action Questionnaire; TMMS: Trait Meta Mood Scale; IRI: Interpersonal Reactivity Index.

Table 6. Comparison of pre-post training mean change by group for self- and observer-reported mindfulness

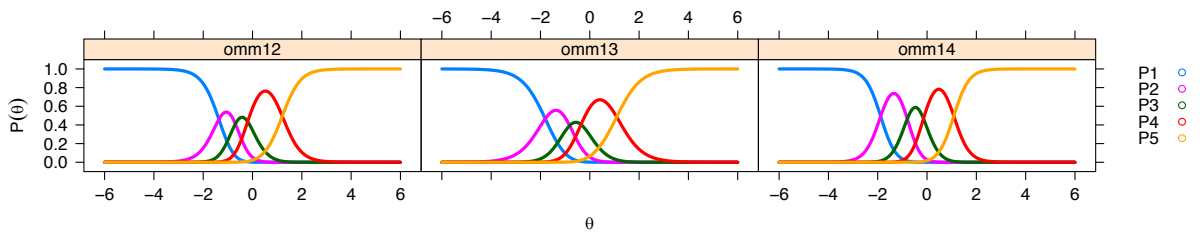
	MT Group n=34		Control Group n=29		Independent sample t-test			SMD	
	M	(SD)	M	(SD)	t	95%CI	p	d	95% CI
Self-reported mindfulness	0.18	(0.74)	-0.10	(0.59)	1.65	-0.06, 0.64	0.10	0.42	-0.10, 0.93
OMM total score	1.62	(3.92)	0.71	(3.56)	0.94	-1.03, 2.86	0.35	0.24	-0.27, 0.75
OMM Awareness	0.97	(1.75)	0.07	(1.68)	2.02	0.01, 1.79	0.05	0.52	0.01, 1.04
OMM Acceptance	0.47	(1.65)	0.39	(2.20)	0.15	-0.92, 1.07	0.88	0.04	-0.47, 0.55
OMM Attentiveness	0.19	(1.69)	0.25	(1.62)	-0.15	-0.92, 0.80	0.89	-0.04	-0.54, 0.47

Sample C data. MT: Mindfulness Training; p: significance of t-test with α=0.05; SMD: Standardised Mean Difference effect estimate; d: Cohen's d; MAAS: Mindful Awareness and Attention Scale (range 1-6); OMM: Observed Mindfulness Measure (range 9-45). OMM Awareness, Acceptance and Attentiveness (range=3-15).

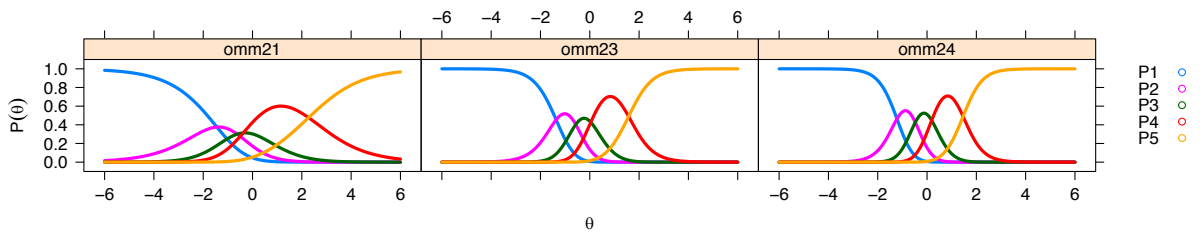
OMM Attentiveness



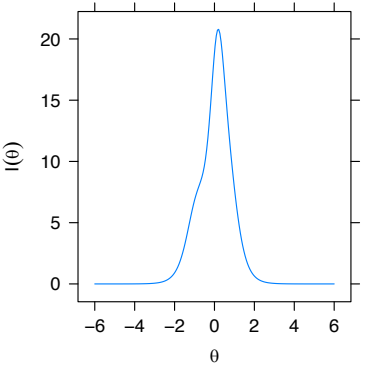
OMM Awareness



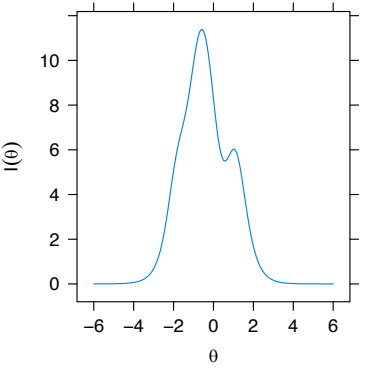
OMM Acceptance



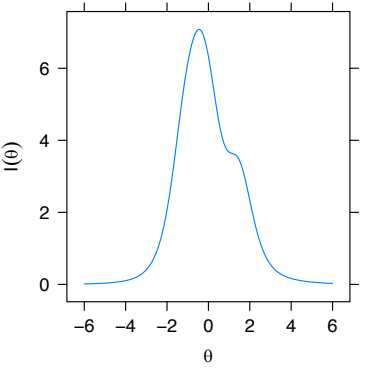
OMM Attentiveness

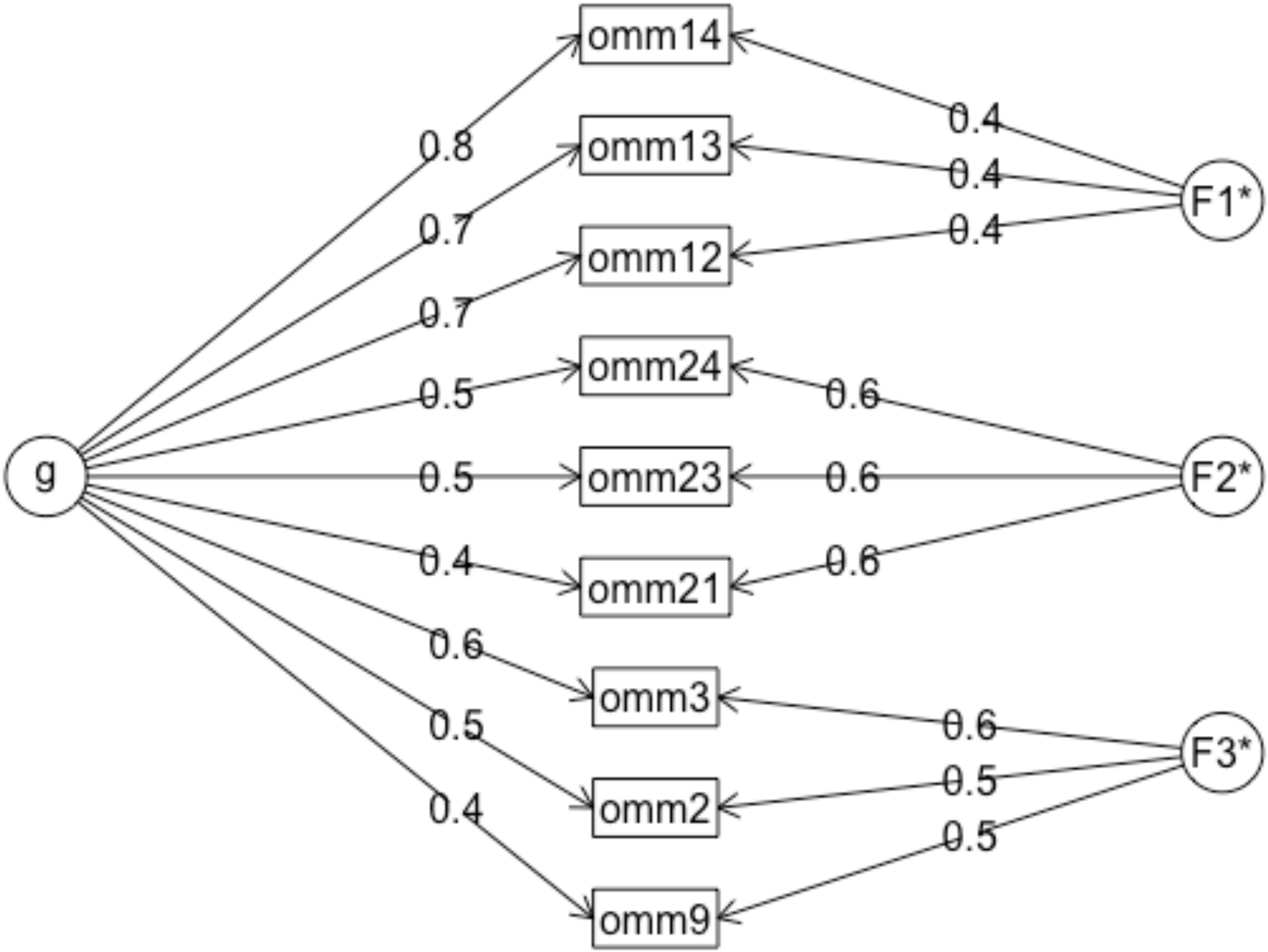


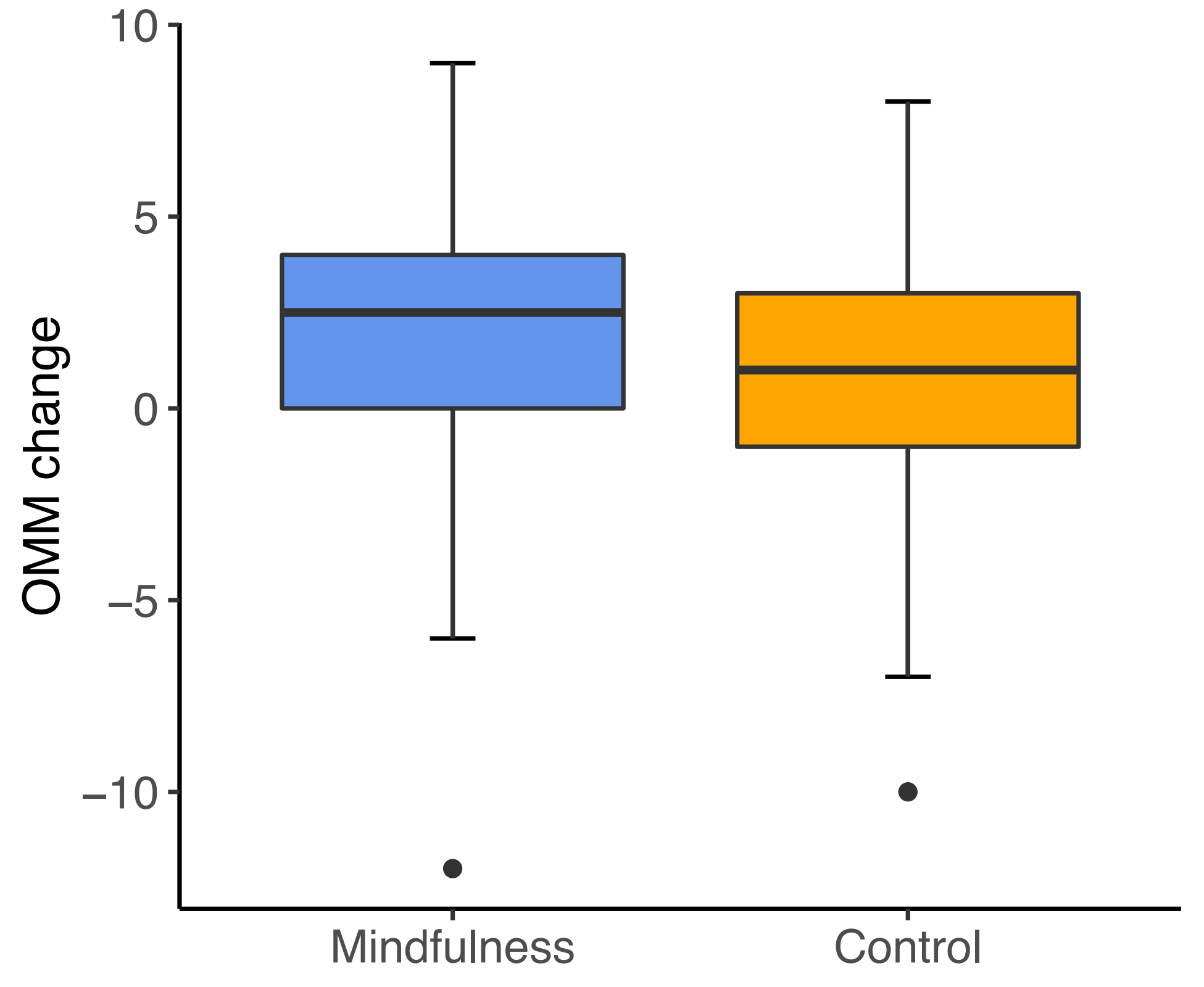
OMM Awareness



OMM Acceptance




omm14
omm13
omm12
omm24
omm23
omm21
omm3
omm2
omm9
g
F1*
F2*
F3*
0.8
0.7
0.7
0.5
0.5
0.4
0.6
0.5
0.4
0.4
0.4
0.4
0.6
0.6
0.6
0.5
0.5


10
5
0
−5
−10
OMM change
Mindfulness
Control

SUPPLEMENTARY INFORMATION

Is mindfulness a noticeable quality? Development and Validation of the Observed Mindfulness Measure

[AUTHOR NAMES TO BE RETURNED PRIOR TO PUBLICATION]

CONTENTS:

		Page
Brief evidence summary: Mechanisms and prosocial outcomes of mindfulness practice		2
Figure S.1.	Mechanisms and outcomes of mindfulness practice	4
Table S.1.	Original self-report items reworded for observer-report with expert adviser recommendations	9
Table S.2.	Survey administration schedule	10
Table S.3.	Item loadings from exploratory factor analysis in Sample A (29 items)	14
Table S.4.	Serial confirmatory factor analyses: influence of items if deleted (18 items)	15
Figure S.2.	Nomological network relationships for the Observed Mindfulness Measure at whole scale and factor level	16

Brief evidence summary: Mechanisms and prosocial outcomes of mindfulness practice

Mindfulness is characterized by intentionally paying attention to current experiences with an open and curious attitude (Bergomi, Tschacher, & Kupper, 2013; Bishop et al., 2004). This natural human quality is consistently associated with positive psychological resources including resilience, optimism, self-efficacy, wellbeing and life satisfaction (Jha, Stanley, Kiyonaga, Wong, & Gelfand, 2010; Khoury, Sharma, Rush, & Fournier, 2015; Roche, Haar, & Luthans, 2014). Research has shown mindfulness can be developed through training programs involving guided learning and meditative practices (Baer, Crane, Miller, & Kuyken, 2019; Kabat-Zinn, 2003).

Outcomes for mindfulness practitioners (internal effects)

Beneficial effects are reported following mindfulness training for a range of stress-related health problems. For example, clinical symptoms improve for people living with anxiety, depression and post-traumatic stress disorder (Boyd, Lanius, & McKinnon, 2018; de Abreu Costa, D'Alò de Oliveira, Tatton-Ramos, Manfro, & Salum, 2018), hypertension and diabetes (Conversano et al., 2021; Ni, Ma, & Li, 2020). Mindfulness training is also frequently used in clinical settings to help reduce the burden of chronic pain (Reiner, Tibi, & Lipsitz, 2013) and is offered as an adjunct treatment to support coping and recovery for people diagnosed with cancer, rheumatoid arthritis, fibromyalgia, irritable bowel syndrome, cardiovascular disease and metabolic disorders (Carlson, 2012). Evidence of improved coping following mindfulness training is supported by a sub-sample of studies showing neuro-biological change in physiological stress and regulatory functioning (Creswell, Lindsay, Villalba, & Chin, 2019; Garland, Hanley, Baker, & Howard, 2017; Lutz et al., 2009). Mindfulness training is a behavioural intervention that is also frequently offered in non-clinical settings, such as in

workplace and community samples, where the people who participate report symptom improvement for chronic stress, depression and anxiety, (Carlson, 2012; Reiner et al., 2013)as well as overall quality of life (Bartlett et al., 2019; Gotink et al., 2015; Khoury et al., 2013).

Outcomes beyond individual practitioners (external effects)

Collectively, there is good empirical evidence that mindfulness training is a positive, skills-based psychological intervention that cultivates protective and beneficial personal resources. For example, emerging research suggests mindfulness training improves interpersonal relationships, empathy and compassion (Blewitt et al., 2015; Condon, Desbordes, Miller, & DeSteno, 2013; Donald et al., 2019; Kreplin, Farias, & Brazil, 2018). Further, a small number of studies have provided some evidence that a person's mindfulness may be associated with leadership, work performance, organisational citizenship and safety behaviours (Good et al., 2016; Hyland, Lee, & Mills, 2015; Reb & Choi, 2014). Familial studies indicate improved intimate partner relationship satisfaction after mindfulness training (Barnes, Brown, Krusemark, Campbell, & Rogge, 2007; Gillespie, Davey, & Flemke, 2015; Karremans, Schellekens, & Kappen, 2017; Khaddouma, Coop Gordon, & Strand, 2017) and that higher parent mindfulness may lead to fewer instances of child aggression and parental discipline (Singh et al., 2010; Siu, Ma, & Chui, 2016). Positive associations have also been found in health care settings between clinician mindfulness and patient-reported quality of care (Cohen-Katz et al., 2005; Horner, Piercy, Eure, & Woodard, 2014; Singh et al., 2015). Further, in education settings, school teachers have reported being more able to reappraise challenging situations following mindfulness training, leading to improved self-care and student engagement (Sharp & Jennings, 2016; Taylor et al., 2016).

Mechanisms of internal and external outcomes

The functional processes involved in mindfulness training are depicted in **Figure 1**.

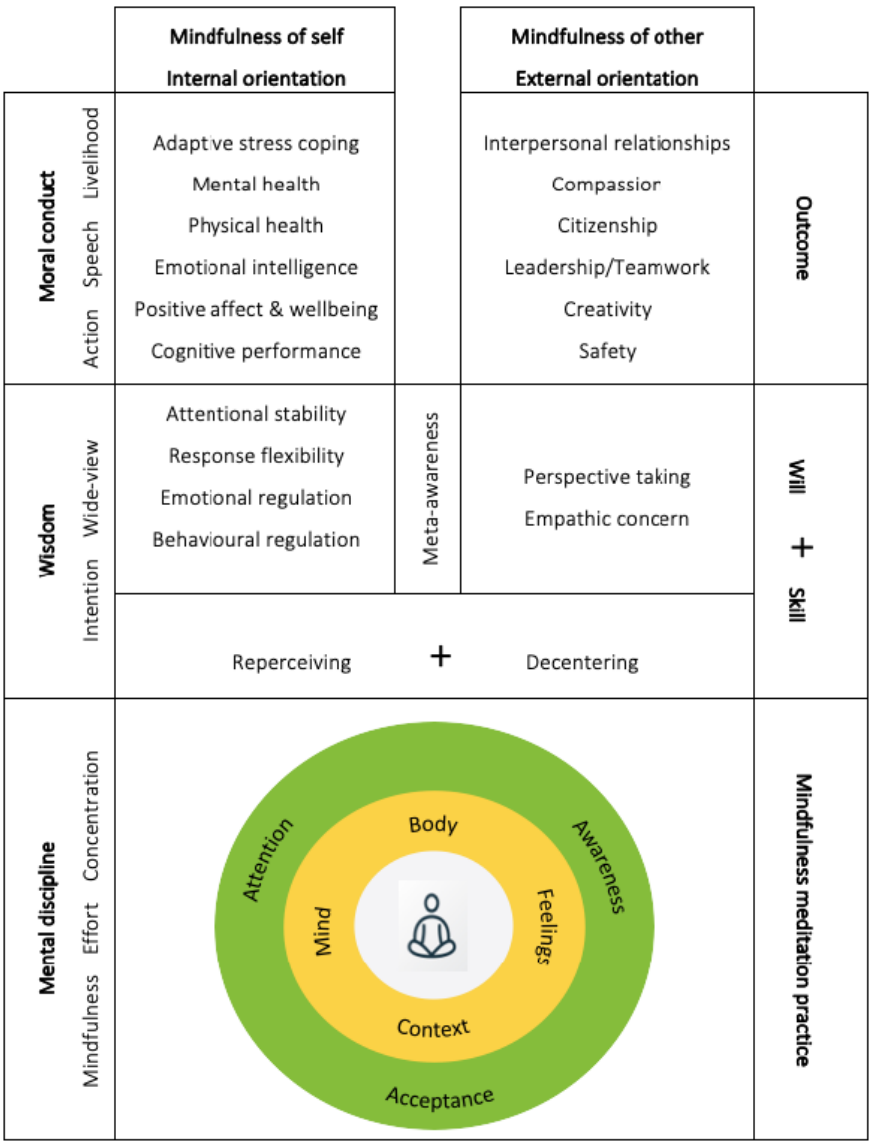


Figure S. 1. Mechanisms and outcomes of mindfulness practice

Mindfulness meditation provides the opportunity to practice the skilful application of mindful awareness, attentiveness and acceptance of internal and external experiences as they occur (Crane et al., 2016). In practice this involves paying attention to the meditation target (flow of breath, body sensations, sounds, thoughts, emotional tones) with an open, non-reactive, non-judging (mindful) attitude. When the attention wanders to thoughts, sensations, emotions and other experiences that arise during the meditation this is noted and accepted as having occurred without mental elaboration, and the attention is returned to the meditation

target (Creswell, 2016; Kabat-Zinn, 2013). Practicing accepting the presence of experiences and detaching from them during meditation develops mental discipline and attentional control (Hölzel et al., 2011). The sustained practice of mindfulness meditation fosters *decentering* and *reperceiving*, which are instrumental in the conscious appraisal of experiences from differing viewpoints or perspectives (Creswell & Lindsay, 2014; Folkman, 2013). These skills are understood to mediate stress reactivity by altering appraisals, increasing self-regulation and enabling response flexibility (Hölzel et al., 2011; Keng, Smoski, & Robins, 2011; Shapiro, Carlson, Astin, & Freedman, 2006). Mindfulness meditation also cultivates *meta-awareness*, which is an important element of psychological flexibility as it enables the practitioner to observe the play of concurrent internal and external information sources that inform situational appraisal, and provides perceptual space with which to choose an appropriate response (Cayoun, 2011; Chambers, Gullone, & Allen, 2009; Hargus, Crane, Barnhofer, & Williams, 2010; Jankowski & Holas, 2014).

Assessing potential external effects of mindfulness

To understand the impact of a person's mindfulness that extend beyond the self, it is vital to first grasp what it is about a person's mindfulness that is noticeable to others. This work can then inform studies into how the internal qualities and personal practices of mindfulness can yield external effects on others. Our paper presents the development of a 9-item questionnaire that measures the extent to which a known other person behaves mindfully: with attentiveness, awareness and acceptance.

Please refer to the full article or contact the corresponding authors [EMAIL ADDRESSES TO BE RETURNED PRIOR TO PUBLICATION] for more information.

REFERENCES

Baer, R.A., Crane, C., Miller, E., and Kuyken, W. (2019). Doing no harm in mindfulness-based programs: Conceptual issues and empirical findings. Clinical Psychology Review.

Barnes, S., Brown, K.W., Krusemark, E., Campbell, W.K., and Rogge, R.D. (2007). The role of mindfulness in romantic relationship satisfaction and responses to relationship stress. Journal of Marital and Family Therapy 33, 482-500.

Bartlett, L., Martin, A., Neil, A.L., Memish, K., Otahal, P., Kilpatrick, M., and Sanderson, K. (2019). A Systematic Review and Meta-Analysis of Workplace Mindfulness Training Randomized Controlled Trials. Journal of Occupational Health Psychology 24, 108-126.

Bergomi, C., Tschacher, W., and Kupper, Z. (2013). The assessment of mindfulness with self-report measures: Existing scales and open issues. Mindfulness 4, 191-202.

Bishop, S.R., Lau, M., Shapiro, S., Carlson, L.E., Anderson, N.D., Carmody, J., Segal, Z.V., Abbey, S., Speca, M., and Velting, D. (2004). Mindfulness: A Proposed Operational Definition. Clinical Psychology: Science and Practice 11, 230-241.

Blewitt, L., Wang, K., Nguyen, H., Johnson, A., Pidial, K., and Yu, N. (2015). Mindfulness: Creating the Space for Compassionate Care. Industrial and Organizational Psychology 8, 706-710.

Boyd, J.E., Lanius, R.A., and Mckinnon, M.C. (2018). Mindfulness-based treatments for posttraumatic stress disorder: a review of the treatment literature and neurobiological evidence. Journal Of Psychiatry & Neuroscience 43, 7-25.

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

Carlson, L.E. (2012). Mindfulness-Based Interventions for Physical Conditions: A Narrative Review Evaluating Levels of Evidence. ISRN Psychiatry.

Cayoun, B.A. (2011). Mindfulness-integrated CBT: Principles and practice. Chichester, UK: John Wiley & Sons.

Chambers, R., Gullone, E., and Allen, N.B. (2009). Mindful emotion regulation: An integrative review. Clinical Psychology Review 29, 560-572.

Cohen-Katz, J., Wiley, S., Capuano, T., Baker, D.M., Deitrick, L., and Shapiro, S. (2005). The effects of mindfulness-based stress reduction on nurse stress and burnout: a qualitative and quantitative study, part III. Holistic Nursing Practice 19, 78-86.

Condon, P., Desbordes, G., Miller, W.B., and Desteno, D. (2013). Meditation Increases Compassionate Responses to Suffering. Psychological Science (0956-7976) 24, 2125-2127.

Conversano, C., Orrù, G., Pozza, A., Miccoli, M., Ciacchini, R., Marchi, L., and Gemignani, A. (2021). Is Mindfulness-Based Stress Reduction Effective for People with Hypertension? A Systematic Review and Meta-Analysis of 30 Years of Evidence. International Journal of Environmental Research and Public Health 18.

Crane, R.S., Brewer, J., Feldman, C., Kabat-Zinn, J., Santorelli, S., Williams, J.M.G., and Kuyken, W. (2016). What defines mindfulness-based programs? The warp and the weft. Psychological Medicine 47, 1-10.

Creswell, J.D. (2016). Mindfulness Interventions. Annual Review of Psychology 68, 491-516.

Creswell, J.D., and Lindsay, E.K. (2014). How Does Mindfulness Training Affect Health? A Mindfulness Stress Buffering Account. Current Directions in Psychological Science 23, 401-407.

Creswell, J.D., Lindsay, E.K., Villalba, D.K., and Chin, B. (2019). Mindfulness Training and Physical Health. Psychosomatic Medicine 81, 224-232.

Davis, M.H. (1983). Measuring individual differences in empathy: Evidence for a multidimensional approach. Journal of Personality and Social Psychology 44, 113.

De Abreu Costa, M., D'alò De Oliveira, G.S., Tatton-Ramos, T., Manfro, G.G., and Salum, G.A. (2018). Anxiety and Stress-Related Disorders and Mindfulness-Based Interventions: a Systematic Review and Multilevel Meta-analysis and Meta-Regression of Multiple Outcomes. Mindfulness.

Donald, J.N., Sahdra, B.K., Van Zanden, B., Johannes Duineveld, J., Atkins, P.W.B., Marshall, S., and Ciarrochi, J. (2019). Does Your Mindfulness Benefit Others? A Systematic Review and Meta-Analysis of the Link Between Mindfulness and Prosocial Behavior. British Journal of Psychology 110, 101-125.

Folkman, S. (2013). "Stress: Appraisal and Coping," in Encyclopedia of Behavioral Medicine, eds. M.D. Gellman & J.R. Turner. (New York: Springer), 1913-1915.

Garland, E.L., Hanley, A.W., Baker, A.K., and Howard, M.O. (2017). Biobehavioral Mechanisms of Mindfulness as a Treatment for Chronic Stress: An RDoC Perspective. Chronic Stress 1, 1–14.

Gillespie, B., Davey, M.P., and Flemke, K. (2015). Intimate Partners' Perspectives on the Relational Effects of Mindfulness-Based Stress Reduction Training: A Qualitative Research Study. Contemporary Family Therapy 37, 396-407.

Good, D.J., Lyddy, C.J., Glomb, T.M., Bono, J.E., Brown, K.W., Duffy, M.K., Baer, R.A., Brewer, J.A., and Lazar, S.W. (2016). Contemplating Mindfulness at Work. Journal of Management 42, 114.

Gotink, R.A., Chu, P., Busschbach, J.J.V., Benson, H., Fricchione, G.L., and Hunink, M.G.M. (2015). Standardised Mindfulness-Based Interventions in Healthcare: An Overview of Systematic Reviews and Meta-Analyses of RCTs. PLoS ONE 10, 1.

Hargus, E., Crane, C., Barnhofer, T., and Williams, J.M.G. (2010). Effects of mindfulness on meta-awareness and specificity of describing prodromal symptoms in suicidal depression. Emotion 10, 34-42.

Hayes, S.C., Luoma, J.B., Bond, F.W., Masuda, A., and Lillis, J. (2006). Acceptance and commitment therapy: Model, processes and outcomes. Behaviour Research and Therapy 44, 1-25.

Hölzel, B.K., Lazar, S.W., Gard, T., Schuman-Olivier, Z., Vago, D.R., and Ott, U. (2011). How Does Mindfulness Meditation Work? Proposing Mechanisms of Action From a Conceptual and Neural Perspective. Perspectives on Psychological Science 6, 537-559.

Horner, J.K., Piercy, B.S., Eure, L., and Woodard, E.K. (2014). A pilot study to evaluate mindfulness as a strategy to improve inpatient nurse and patient experiences. Applied Nursing Research 27, 198.

Hyland, P.K., Lee, R.A., and Mills, M.J. (2015). Mindfulness at Work: A New Approach to Improving Individual and Organizational Performance. Industrial and Organizational Psychology 8, 576-602.

Jankowski, T., and Holas, P. (2014). Review: Metacognitive model of mindfulness. Consciousness and Cognition 28, 64-80.

Jha, A.P., Stanley, E.A., Kiyonaga, A., Wong, L., and Gelfand, L. (2010). Examining the protective effects of mindfulness training on working memory capacity and affective experience. Emotion 10, 54-64.

Kabat-Zinn, J. (2003). Mindfulness-Based Interventions in Context: Past, Present, and Future. Clinical psychology: Science and practice 10, 144-156.

Kabat-Zinn, J. (2013). Full Catastrophe Living: How to cope with stress, pain and illness using mindfulness meditation. London: Piatkus.

Karremans, J.C., Schellekens, M.P., and Kappen, G. (2017). Bridging the Sciences of Mindfulness and Romantic Relationships. Personality and Social Psychology Review 21, 29-49.

Keng, S.-L., Smoski, M.J., and Robins, C.J. (2011). Effects of mindfulness on psychological health: A review of empirical studies. Clinical Psychology Review 31, 1041-1056.

Khaddouma, A., Coop Gordon, K., and Strand, E.B. (2017). Mindful Mates: A Pilot Study of the Relational Effects of Mindfulness-Based Stress Reduction on Participants and Their Partners. Family Process 56, 636-651.

Khoury, B., Lecomte, T., Fortin, G., Masse, M., Therien, P., Bouchard, V., Chapleau, M.A., Paquin, K., and Hofmann, S.G. (2013). Mindfulness-based therapy: a comprehensive meta-analysis. Clinical Psychology Review 33, 763-771.

Khoury, B., Sharma, M., Rush, S.E., and Fournier, C. (2015). Mindfulness-based stress reduction for healthy individuals: A meta-analysis. Journal of Psychosomatic Research 78, 519-528.

Kreplin, U., Farias, M., and Brazil, I.A. (2018). The limited prosocial effects of meditation: A systematic review and meta-analysis. Scientific Reports 8, 2403.

Lutz, A., Slagter, H.A., Rawlings, N.B., Francis, A.D., Greischar, L.L., and Davidson, R.J. (2009). Mental training enhances attentional stability: neural and behavioral evidence. Journal of Neuroscience 29, 13418-13427.

Ni, Y., Ma, L., and Li, J. (2020). Effects of Mindfulness-Based Stress Reduction and Mindfulness-Based Cognitive Therapy in People With Diabetes: A Systematic Review and Meta-Analysis. J Nurs Scholarsh.

Reb, J., and Choi, E. (2014). "Mindfulness in organizations," in Psychology of Meditation.: Nova Science Publishers, Hauppauge, NY), 279-309.

Reiner, K., Tibi, L., and Lipsitz, J.D. (2013). Do Mindfulness-Based Interventions Reduce Pain Intensity? A Critical Review of the Literature. Pain Medicine 14, 230-242.

Roche, M., Haar, J.M., and Luthans, F. (2014). The role of mindfulness and psychological capital on the well-being of leaders. Journal of Occupational Health Psychology 19, 476.

Salovey, P., Mayer, J.D., Goldman, S.L., Turvey, C., and Palfai, T.P. (1995). "Emotional attention, clarity, and repair: Exploring emotional intelligence using the Trait Meta-Mood Scale," in Emotion, Disclosure, and Health, ed. J.W. Pennebaker. (Washington DC: American Psychological Association), 125-154.

Shapiro, S.L., Carlson, L.E., Astin, J.A., and Freedman, B. (2006). Mechanisms of mindfulness. Journal Of Clinical Psychology 62, 373-386.

Sharp, J.E., and Jennings, P.A. (2016). Strengthening teacher presence through mindfulness: What educators say about the Cultivating Awareness and Resilience in Education (CARE) program. Mindfulness 7, 209-218.

Singh, N.N., Lancioni, G.E., Karazsia, B.T., Myers, R.E., Winton, A.S.W., Latham, L.L., and Nugent, K. (2015). Effects of training staff in MBPBS on the use of physical restraints, staff stress and turnover, staff and peer injuries, and cost effectiveness in developmental disabilities. Mindfulness 6, 926-937.

Singh, N.N., Lancioni, G.E., Winton, A.S.W., Singh, J., Singh, A.N., Adkins, A.D., and Wahler, R.G. (2010). Training in Mindful Caregiving Transfers to Parent–Child Interactions. Journal of Child and Family Studies 19, 167-174.

Siu, A.F.Y., Ma, Y., and Chui, F.W.Y. (2016). Maternal Mindfulness and Child Social Behavior: the Mediating Role of the Mother-Child Relationship. Mindfulness, 1-7.

Taylor, C., Harrison, J., Haimovitz, K., Oberle, E., Thomson, K., Schonert-Reichl, K., and Roeser, R.W. (2016). Examining ways that a mindfulness-based intervention reduces stress in public school teachers: A mixed-methods study. Mindfulness 7, 115-129.

Table S.1. Original self-report items reworded for observer-report with expert adviser recommendations

Origin of item	Original self-report item	Reworded item for observer respondent	Expert reviewer comment ✓= include; x = exclude; ? = unsure					Include for testing?
MAAS2	I break or spill things because of carelessness, not paying attention, or thinking of something else	The person often breaks or spills things due to carelessness or not paying attention	✓	-	✓	✓	✓	✓
MAAS3	I find it difficult to stay focused on what's happening in the present	The person has difficulty staying focused on what is happening to/around them as it occurs	?	-	✓	✓	✓	✓
MAAS4	I tend to walk quickly to where I'm going without paying attention to what I experience along the way	The person tends to walk quickly to their destination without focusing on what he or she experiences along the way	x	-	✓	x	x	x
MAAS6	I forget a person's name almost as soon as I've been told it for the first time	The person will quickly forget the name of someone he or she has met for the first time	?	-	✓	x	✓	x
MAAS7	It seems I am 'running on automatic', without much awareness of what I'm doing	The person seems to 'run on automatic' without much awareness of what he or she is doing	?	-	✓	✓	✓	✓
MAAS8	I rush through activities without being really attentive to them	The person rushes through activities without really being attentive to them	✓	-	✓	✓	x	x
MAAS9	I get so focused on the goal I want to achieve that I lost touch with what I'm doing right now to get there	The person is very goal oriented and can lose touch with what he or she is doing to achieve it	✓	x	?	✓	✓	x
MAAS11	I find myself listening to someone with one ear, doing something else at the same time	The person does other things while listening to someone	✓	-	✓	x	✓	✓
MAAS13	I find myself preoccupied with the future or the past	The person seems preoccupied with the future or the past	✓	-	✓	x	✓	✓
KIMS2	I'm good at finding the words to describe my feelings	The person is good at finding the words to describe his or her feelings	✓	-	✓	x	✓	✓
KIMS6	I can easily put my beliefs, opinions and expectations into words	The person finds it easy to express his or her beliefs, opinions and expectations	✓	-	✓	x	✓	✓
KIMS7	When I'm doing something, I'm only focused on what I'm doing, nothing else	When doing something, the person focuses only on that one thing, nothing else	✓	-	?	x	✓	x

Table S.1. Original self-report items reworded for observer-report with expert adviser recommendations

Origin of item	Original self-report item	Reworded item for observer respondent	Expert reviewer comment ✓= include; x = exclude; ? = unsure					Include for testing?
KIMS8	I tend to evaluate whether my perceptions are right or wrong	The person tends to evaluate his or her own perceptions as right or wrong	?	-	✓	x	✓	x
KIMS10	I'm good at thinking of words to express my perceptions, such as how things taste, smell, or sound	The person is good at thinking of words to express his or her perceptions, such as how things taste, smell or sound	✓	-	✓	x	✓	✓
KIMS14	It's hard for me to find the words to describe what I'm thinking	It is hard for this person to find the words to describe what he or she is thinking	?	-	✓	x	✓	x
KIMS18	I have trouble thinking of the right words to express how I feel about things	The person has trouble finding words to express how he or she feels about things	-	-	✓	x	x	x
KIMS19	When I do things, I get totally wrapped up in them and don't think about anything else	This person gets totally wrapped up in what he or she is doing, and doesn't think of anything else	✓	-	?	x	x	x
KIMS22	When I have a sensation in my body, it's difficult for me to describe it because I can't find the right words	It is difficult for this person to describe physical sensations because he or she can't find the right words	?	-	✓	x	✓	x
KIMS23	I don't pay attention to what I'm doing because I'm daydreaming, worrying or otherwise distracted	The person doesn't pay attention to what he or she is doing, because of daydreaming, worrying or other distractions	✓	-	✓	✓	✓	✓
KIMS24	I tend to make judgments about how worthwhile or worthless my experiences are	The person tends to make judgments about how worthy his or her experiences are	-	-	✓	✓	✓	x
KIMS26	Even when I'm terribly upset, I can find a way to put it into words	Even when terribly upset, this person can find words to explain what is going on	✓	-	✓	x	✓	✓
KIMS31	I tend to do several things at once rather than focusing on one thing at a time	The person tends to do several things at once rather than focusing on one thing at a time	✓	-	✓	x	✓	x
KIMS34	My natural tendency is to put my experiences into words	The person has a natural tendency to put his or her experiences into words	✓	-	✓	x	✓	x

Table S.1. Original self-report items reworded for observer-report with expert adviser recommendations

Origin of item	Original self-report item	Reworded item for observer respondent	Expert reviewer comment ✓= include; x = exclude; ? = unsure					Include for testing?
KIMS37	I pay attention to how my emotions affect my thoughts and behaviour	The person seems aware of how emotions affect his or her thoughts and behaviour	?	-	✓	✓	x	✓
KIMS38	I get completely absorbed in what I'm doing, so that all my attention is focused on it	The person gets completely absorbed in what he or she is doing, so that all their attention is focused on it	✓	-	?	x	✓	x
PHLMS11	When someone asks how I am feeling, I can identify my emotions easily	When asked how he or she is feeling, the person can identify their emotions easily	✓	-	✓	✓	x	✓
PHLMS19	When talking with other people, I am aware of the emotions I am experiencing	The person seems aware of his or her own emotions when iteracting with others	✓	-	✓	✓	✓	✓
LMS1	I like to investigate things	The person likes to investigate things	✓	-	✓	✓	✓	✓
LMS2	I generate few novel ideas	The person generates few novel ideas	✓	-	✓	x	✓	x
LMS3	I make many novel contributions	The person makes many novel contributions	✓	-	✓	x	✓	x
LMS4	I seldom notice what other people are up to	The person seldom notices what other people are up to	✓	-	✓	✓	✓	✓
LMS5	I avoid thought provoking conversations	The person avoids thought provoking conversations	?	-	✓	✓	✓	✓
LMS6	I am very creative	The person is very creative	✓	-	?	x	✓	x
LMS7	I am very curious	The person is curious	✓	-	✓	x	✓	x
LMS8	I try to think of new ways of doing things	The person tries to think of new ways of doing things	✓	-	✓	x	✓	✓
LMS9	I am rarely aware of changes	The person is rarely aware of changes	?	-	✓	✓	x	x
LMS10	I like to be challenged intellectually	The person likes to be challenged intellectually	?	-	✓	✓	✓	x
LMS11	I find it easy to create new and effective ideas	The person finds it easy to create new and effective ideas	✓	-	✓	x	x	x
LMS12	I am rarely alert to new developments	The person is rarely alert to new developments	✓	-	✓	✓	✓	✓
LMS13	I like to figure out how things work	The person likes to figure out how things work	✓	-	✓	x	✓	x
LMS14	I am not an original thinker	The person is not an original thinker	✓	-	?	x	✓	x
FMI4	I am able to appreciate myself	The person appears to appreciate his or her self	?	-	✓	✓	✓	✓

Table S.1. Original self-report items reworded for observer-report with expert adviser recommendations

Origin of item	Original self-report item	Reworded item for observer respondent	Expert reviewer comment ✓= include; x = exclude; ? = unsure					Include for testing?
FMI5	I pay attention to what is behind my actions	The person pays attention to what is behind his or her actions	?	-	✓	✓	x	x
FMI6	I see my mistakes and difficulties without judging them	The person appears to accept his or her mistakes and difficulties without judging them	?	-	✓	✓	✓	✓
FMI8	I accept unpleasant experiences	The person appears to recover well from unpleasant or stressful experiences	?	-	✓	✓	✓	✓
FMI9	I am friendly to myself when things go wrong	The person does not appear to get upset with him- or herself when things go wrong	✓	-	✓	✓	✓	✓
FMI11	In difficult situations I can pause without immediately reacting	The person can pause before reacting to difficult situations	✓	-	✓	✓	✓	✓
FMI12	I experience moments of inner peace, even when things get hectic and stressful	The person remains calm, even when things get hectic and stressful	✓	-	✓	✓	✓	✓
FMI13	I am impatient with myself and others	The person is impatient with him or herself and with others	✓	-	✓	✓	✓	✓
FMI14	I am able to smile when I notice how sometimes I make life difficult	Sometimes this person makes life difficult, but is able to notice and smile about it	✓	-	✓	✓	x	x

MAAS: Mindful Attention and Awareness Scale; KIMS: Kentucky Inventory of Mindfulness Skills; PHLMS: Philadelphia Mindfulness Scale; LMS: Langer Mindfulness Scale; FMI: Freiburg Mindfulness Inventory

Table S.2. Survey administration schedule

	Sample A	Sample B				Sample C			
Respondents	Observers	Participants		Observers		Participants		Observers	
Timepoint	Once only	T0	T1	T0	T1	Pre-	Post-	Pre-	Post-
Survey instruments									
Demographics	✓	✓		✓		✓		✓	
Potential items retained after Stage 1	✓								
Potential items retained after Stage 2				✓	✓			✓	✓
Potential items retained after Stage 2 (in original self-report format)		✓	✓			✓	✓		
Mindfulness (MAAS)		✓	✓			✓	✓		
Emotional intelligence (TMMS)		✓							
Acceptance (AAQ)		✓							
Empathy (IRI)		✓							

Sample A completed a single instance survey; Sample B completed two surveys, T0 denotes the first instance, T1 denotes seven days after T0; Sample C completed two surveys, Pre-intervention and Post-intervention. OMM: Observed Mindfulness Measure; OMM-C: Observed Mindfulness Measure items in self-report format; MAAS: Mindful Awareness and Attention Scale; TMMS: Trait Meta Mood Scale; AAQ: Acceptance and Action Questionnaire; IRI: Interpersonal Reactivity Index

Table S.3. Item loadings from exploratory factor analysis in Sample A (29 items)

OMM29 items		Three-factor model			Single factor model
		F1	F2	F3	
OMM1	The person has many mishaps or accidents	0.02	-0.03	0.44	0.29
*OMM2	The person has difficulty staying focused on what is happening to/around them as it occurs	0.09	-0.05	0.72	0.51
*OMM3	The person seems to 'run on automatic' without much awareness of what he or she is doing	0.18	-0.14	0.74	0.54
*OMM4	The person does other things while listening to someone	-0.03	-0.03	0.51	0.29
OMM5	The person seems preoccupied with the future or the past	-0.08	0.33	0.40	0.44
OMM6	The person is good at finding the words to describe his or her feelings	0.75	-0.02	-0.03	0.63
*OMM7	The person can express his or her beliefs, opinions and expectations	0.75	-0.06	-0.14	0.52
*OMM8	The person can describe how things taste, smell or sound	0.70	-0.21	-0.07	0.41
*OMM9	The person doesn't pay attention to what he or she is doing	0.02	-0.11	0.78	0.45
*OMM11	Even when terribly upset, this person can find words to explain what is going on	0.64	0.12	0.05	0.70
*OMM12	The person seems aware of how emotions affect his or her thoughts and behaviour	0.63	0.13	0.07	0.71
*OMM13	When asked how he or she is feeling, the person can identify their emotions easily	0.82	-0.04	0.00	0.70
*OMM14	The person seems aware of his or her own emotions when interacting with others	0.67	0.12	0.00	0.69
*OMM15	The person notices what other people are up to	0.66	-0.16	0.14	0.57
*OMM16	The person avoids thought provoking conversations	0.42	-0.08	0.22	0.46
OMM17	The person tries to think of new ways of doing things	0.36	0.38	-0.09	0.58
OMM18	The person is alert to new developments	0.42	0.29	-0.03	0.60
*OMM19	The person appears to appreciate him- or her self	0.19	0.56	-0.15	0.50
*OMM20	The person appears to accept his or her mistakes and difficulties without judging them	0.03	0.83	-0.16	0.54
*OMM21	The person seems to recover well from unpleasant or stressful experiences	0.02	0.88	-0.18	0.56
OMM22	The person gets upset with him- or herself when things go wrong	-0.15	0.44	0.03	0.22
*OMM23	The person can pause before reacting to difficult situations	0.09	0.51	0.13	0.57
*OMM24	The person remains calm, even when things get hectic and stressful	0.00	0.57	0.11	0.52
OMM25	The person is impatient with him or herself and with others	-0.16	0.22	0.31	0.24
OMM26	The person pays attention to what I am saying when we are talking	0.43	0.09	0.26	0.64
OMM27	The person welcomes new ideas	0.43	0.42	-0.10	0.66
OMM28	The person becomes upset when faced with changing circumstances	-0.06	0.34	0.28	0.41
OMM29	The person avoids difficult situations	0.18	0.12	0.22	0.41
*OMM30	The person holds onto past experiences	-0.12	0.55	0.21	0.46

* denotes items that met cut points for retention. OMM10 was excluded based on interpretability in Stage 1.

Table S.4. Serial confirmatory factor analyses: influence of items if deleted (18 items)

Sample	Model	Excluded items	RMSEA	$\chi^2 < 0.05$	CFI	TLI	MI > 10 (no. items)
A (n=307)	All items	-	0.082	✓	0.885	0.866	15
B+C (N=215)	Model 1	-	0.074	✓	0.902	0.886	10
	Individual items excluded based on potential redundancy/duplication						
	Model 2	omm7	0.073	✓	0.904	0.888	10
	Model 3	omm19	0.077	✓	0.903	0.886	11
	Model 4	omm23	0.069	✓	0.917	0.902	5
	Model 5	omm12	0.073	✓	0.904	0.888	9
	Model 6	omm20	0.071	✓	0.913	0.898	4
	Model 7	omm21	0.075	✓	0.901	0.883	9
	Model 8	omm24	0.07	✓	0.914	0.899	8
	Model 9	omm14	0.074	✓	0.900	0.883	11
	Model 10	omm6	0.074	✓	0.909	0.893	11
	Individual items excluded based on discrimination/difficulty scores						
	Model 11	omm4	0.079	✓	0.899	0.882	11
	Model 12	omm8	0.070	✓	0.917	0.903	11
	Model 13	omm11	0.073	✓	0.908	0.893	15
	Model 14	omm15	0.077	✓	0.902	0.885	12
	Model 15	omm19	0.077	✓	0.903	0.886	11
	Model 16	omm30	0.076	✓	0.904	0.888	11
	Combination models						
	Model 17	Models 11:16	0.085	✓	0.929	0.908	10
	Model 18	Mod 17 + omm7	0.069	✓	0.959	0.946	6
	Model 19	Model 18 + omm23	0.055	✓	0.976	0.966	3
	Model 20	Model 18 + omm21	0.062	✓	0.970	0.958	5
	Model 21	Model 18 + omm20	0.037	0.125	0.989	0.985	1
	Model 22	Model 21 + omm6	0.019	0.364	0.998	0.997	0

RMSEA: Root Mean Square of Error Approximation; χ^2: Degrees of freedom for the Model Fit Test Statistic using Maximum Likelihood estimator for a three-factor structure; CFI: Comparative Fit Index; TLI: Tucker Lewis Index; MI: Modification Indices

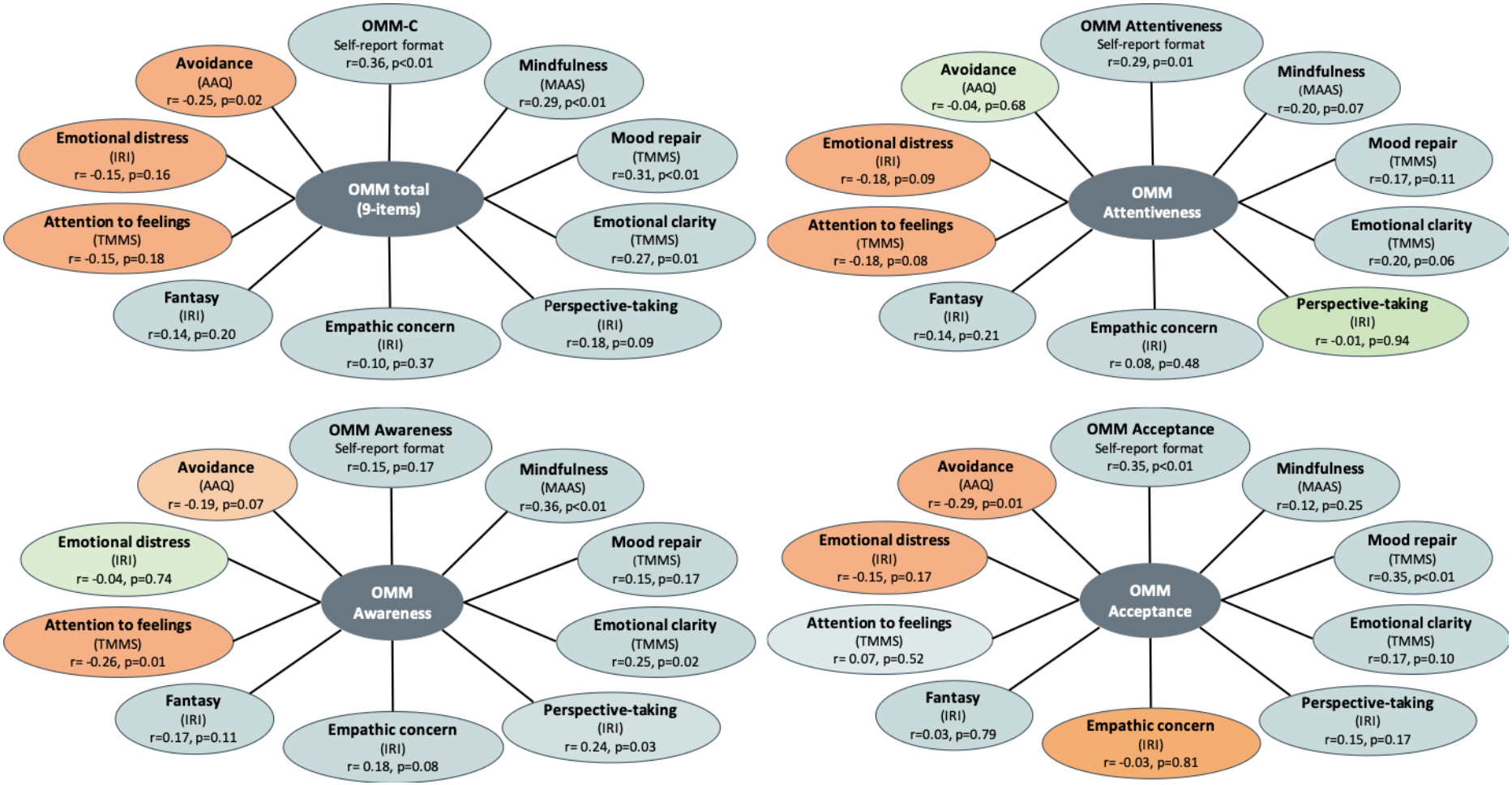


Fig S.2. Nomological network relationships for the Observed Mindfulness Measure at whole scale and factor level

OMM: Observed Mindfulness Measure; MAAS: Mindful Attention and Awareness Scale (Brown & Ryan, 2003); TMMS: Trait Meta Mood Scale (Salovey, Mayer, Goldman, Turvey, & Palfai, 1995); IRI: Interpersonal Reactivity Index (Davis, 1983); AAQ: Acceptance and Action Questionnaire (Hayes, Luoma, Bond, Masuda, & Lillis, 2006)