

Development and Initial Validation of the State Four Facet Mindfulness Questionnaire

Assessment

1–29

© The Author(s) 2025



Article reuse guidelines:

sagepub.com/journals-permissions

DOI: 10.1177/10731911251330092

journals.sagepub.com/home/asm



Stephen Raynes¹  and Karen Dobkins¹

Abstract

The current research aimed to provide initial psychometric validation of a new multifaceted mindfulness questionnaire (referred to as the State Four Facet Mindfulness Questionnaire, or the “state-4FMQ” for short) adapted from the commonly used Five Facet Mindfulness Questionnaire (referred to as the “trait-FFMQ”). The research was divided into two pre-registered studies. In both, undergraduates partook in a 20-minute mindfulness meditation (via audio recording), and then answered questions, including the state-4FMQ, pertaining to their experience during the meditation. In Study 2, participants additionally partook in a 20-minute control condition. The state-4FMQ was developed using exploratory factor analysis (EFA; Study 1) and confirmatory factor analysis (CFA; Study 2). In Study 2, a short-form of the state-4FMQ was established, and several additional forms of measurement validity were tested. EFA and CFA results supported a four-factor structure, which was identical to the trait-FFMQ with the exclusion of Nonreactivity. This newly created state-4FMQ, and its short-form, showed good internal consistency as well as convergent, predictive, and construct validity. In addition, it was found that some facets, more than others, predicted momentary well-being. The validity of the state-4FMQ shows that it can be used to measure multiple facets of state mindfulness across a variety of situations.

Keywords

mindfulness, assessment, psychometric properties, State Four Facet Mindfulness Questionnaire (state-4FMQ), validation

Introduction

Mindfulness originates in various ancient spiritual traditions and is most clearly articulated through Buddhist scholarship (Keng et al., 2011). Despite its extensive history, the systematic investigation of mindfulness in scientific contexts has only recently blossomed (see Van Dam et al., 2018). Much of this can be traced to the pioneering work of John Kabat-Zinn in the 1970s, who explored the use of mindfulness meditation in treating patients with chronic pain through an intervention known as Mindfulness-Based Stress Reduction (MBSR; Kabat-Zinn, 1982). Since then, empirical research has demonstrated the beneficial effects of mindfulness on well-being via three main domains (outlined in a literature review by Keng et al., 2011): First are “real-world” interventional studies, which demonstrate the benefits of mindfulness-based interventions (including, but not limited to, MBSR). These interventions typically consist of several weeks of a structured program, designed for clinical populations (see Creswell, 2017; Dawson et al., 2020; Dhillon et al., 2017; Goldberg et al., 2018, 2022; Goyal et al., 2014, p. 201 for reviews). Second are lab-based studies, which demonstrate the benefits of

procedures that employ short-term mindful inductions including, but not limited to, a single meditation (e.g., Erisman & Roemer, 2011; Feldman et al., 2010; Heppner & Shirk, 2018; Johnson et al., 2013, p. 20; Mrazek et al., 2012; Thompson & Waltz, 2007; Zeidan et al., 2015). Third are correlational studies, which demonstrate associations between mindfulness measures (most of which use *trait* mindfulness) and other measures (e.g., Brown et al., 2007; Carpenter et al., 2019; Chu & Mak, 2020; Donald et al., 2019, 2020; Prieto-Fidalgo et al., 2021). Given the frequency by which research examines the advantages and associations of mindfulness, it is critical for the field to operationalize mindfulness with valid and reliable measurement scales. While most of the groundwork has been done with regard to trait mindfulness, the goal of the current

¹Department of Psychology, University of California, San Diego, La Jolla, CA, USA

Corresponding Author:

Karen Dobkins, Department of Psychology, University of California, San Diego, 9500 Gilman Dr, La Jolla, CA 92093, USA.
Email: kdobkins@ucsd.edu

research is to develop a new state mindfulness measure, one that is both multi-dimensional and versatile.

Before arguing the need for a new state mindfulness measure, it is perhaps fruitful to begin by mentioning that early creations of mindfulness measures were trait-based, with the idea that mindfulness is a characteristic that can be developed through practice (noting that others have argued that mindfulness can be as conceptualized as an inherent disposition or a skill (e.g., Burzler & Tran, 2022), a process (e.g., Erisman & Roemer, 2011), an action (e.g., Preissner et al., 2024), and an outcome (Medvedev et al., 2022), all of which have been argued to be distinct from trait mindfulness). How best to define and operationalize mindfulness proved to be challenging, however, due partially to the fact that it derives from various Buddhist texts and traditions that inconsistently, and often vaguely, describe mindfulness in disparate contexts (Gethin, 2011; Grossman, 2008, 2011; Grossman & Van Dam, 2011). For instance, Medvedev (2022) highlights seven different definitions of mindfulness commonly cited in the psychology literature, each with a unique emphasis (for similar discussions see Davidson & Kaszniak, 2015; Lustig et al., 2024; Nilsson & Kazemi, 2016). Based on the myriad of ways mindfulness can be conceptualized, it is perhaps not surprising that there exist many different trait mindfulness measures in the literature (see Bergomi et al., 2013; Sauer et al., 2012 for reviews), and that the correlation between them is often weak or non-existent (e.g., Park et al., 2013; however, see Buhk et al., 2023 for a more recent study showing stronger convergence between trait measures).

These discrepancies also bleed over into challenges in determining whether mindfulness should be considered, and hence operationalized, as a uni-dimensional versus multi-dimensional phenomenon, and if a multidimensional phenomenon, what those dimensions should be (see Quaglia et al., 2015 for detailed discussion). The two most commonly used trait measures of mindfulness exemplify this distinction. While the trait Mindful Attention and Awareness Scale (trait-MAAS; Brown & Ryan, 2003) conceptualizes mindfulness as unidimensional (consisting of attention/awareness), the Five Facet Mindfulness Questionnaire (hereby referred to as the “trait-FFMQ”; Baer et al., 2006) includes the following five dimensions, which we return to later: *Acting with Awareness* (“*ActAware*”; attending to the present moment, as opposed to focusing attention elsewhere or behaving automatically), *Describing* (the ability to express one’s experiences in words), *Nonjudging* (the acceptance of one’s thoughts and emotions without evaluation), *Nonreactivity* (the ability to allow thoughts and emotions to come and go without becoming attached or carried away with them), and *Observing* (attending to or

noticing both internal and external experiences, such as thoughts, emotions, bodily sensations, smells and sounds). Note, however, that the extent to which trait Observing loads onto the superordinate mindfulness construct appears to vary with meditation experience (see Burzler & Tran, 2022).

Although trait measures of mindfulness are commonly used in research studies, there is good reason to develop *state* measures since mindfulness is often referred to—by Buddhists and researchers alike—as a *momentary* state. For example, Jon Kabat-Zinn states that “we are all mindful to one degree or another, moment by moment” (Kabat-Zinn, 2003, p. 145–146). In a similar vein, mindfulness has been described as an inherent and universal human capacity (Brown & Ryan, 2004) that has the potential to improve well-being when experienced (Ludwig & Kabat-Zinn, 2008), as well as a state of consciousness, the qualities of which can vary considerably depending on the context (Brown & Ryan, 2003). In fact, Bishop et al. (2004) influentially proposed that mindfulness is *best* defined as a state-like quality that only exists when one’s attention to their experience is purposely cultivated with an open and non-judgmental attitude. As such, experiencing a state of mindfulness can happen during formal practices of mindfulness, such as meditation, or during instances of daily life (Bishop et al., 2004; Brown & Ryan, 2003, 2004).

In addition to its conceptual validity, one practical advantage of defining mindfulness as a state is that it can then be studied in more diverse and comprehensive ways. For example, state mindfulness can be measured in reference to a *single* experimental manipulation setup in a laboratory (e.g., a single session of meditation, or any other experimental manipulation), or in reference to *multiple* naturally occurring experiences as part of one’s daily life (outside a laboratory). This latter approach, referred to as intensive longitudinal design (ILD), can afford abundant statistical power since multiple time-points (i.e., repeated measures) are obtained from a single participant. Two commonly used ILD approaches are the Experience Sampling Method (ESM) where participants respond to questions in reference to the present or immediately preceding moment through notifications sent at semi-random intervals through a mobile device, and the Day Reconstruction Method (Kahneman et al., 2004) where participants systematically reconstruct the previous day’s events and then respond to questions in reference to each event. These methods allow a much more fine-grained and ecologically valid investigation of the relationship between mindfulness and other psychological constructs (e.g., affect) as compared to the use of a single trait measure of mindfulness. In sum, based on the conceptual validity, and empirical usefulness, of

Table 1. Validated State Mindfulness Measures.

Measure Name	Dimensions	Theoretical Basis	Development Article	Citations ^a
Mindful Awareness and Attention Scale (state-MAAS)	1. Total score (5 items)	Mix	Brown & Ryan, 2003, Study 4	18,606 ^b
Toronto Mindfulness Scale (TMS)	1. Curiosity (6 items) 2. Decentering (7 items)	Mostly Bishop et al.'s 2004 mindfulness definition	Lau et al., 2006	2,034
State Mindfulness Scale (SMS)	1. Total score 2. Mind (15 items) 3. Body (6 items)	Bishop et al.'s 2004 mindfulness definition and Buddhist texts	Tanay & Bernstein, 2013	572
Multidimensional State Mindfulness Questionnaire (MSMQ)	1. Present-moment ... attention (3 items) 2. Acting with awareness (3 items) 3. Nonjudgmental acceptance (3 items)	Mix of the most commonly used mindfulness conceptualizations	Blanke & Brose, 2017	71

Note. Close adaptations of these scales, such as the State Mindfulness Scale for Physical Activity (SMS-PA; Ullrich-French et al., 2017), are not reviewed.

^aGoogle Scholar citations as of February 7, 2024. ^bThis includes citations for the trait-MAAS, which was developed in the same article and thus cannot be separated.

considering mindfulness a momentary state, it behooves the field of mindfulness research to have readily available state measures of mindfulness.

To date, there exist four validated state measures of mindfulness, which to our knowledge have yet to be reviewed altogether. In the next sections, we review these scales (in descending order of a number of citations), by addressing (a) the framework used to guide the creation of the scale, (b) the “referent experience” the scale was tested with (e.g., a single experimental manipulation vs. an ILD), (c) the multi-dimensionality of the scale, and (d) how the scale was validated (summarized in Table 1). This is then followed by a section arguing the need for a new state measure fashioned after the widely used trait-FFMQ, which is the goal of the current study.

Of all the state mindfulness measures, the most used in research is the five-item State Mindful Attention and Awareness Scale (state-MAAS; Brown & Ryan, 2003). This scale was fashioned after the trait MAAS (15 items), and thus the two scales (state and trait) have analogous constructs. The MAAS is theoretically grounded in a broad and vague mix of sources including “our personal experience and knowledge of mindfulness (and mindlessness), published writings on mindfulness and attention, and existing scales assessing conscious states of various kinds.” (Brown & Ryan 2003, p. 825). The scale was purposely designed to capture what the authors argue is the one central component of mindfulness—attention to, and awareness of, the present moment—rather than any other components commonly associated with mindfulness such as an accepting attitude. Accordingly, the state-MAAS was developed with one dimension, “Attention and Awareness” (e.g., “I was doing something automatically, without being aware of what I was doing”, reverse scored), producing a total score across the five reverse-scored items. The scale was developed using an ESM design (92 students recorded three times per day for 14 consecutive days), though the scale has since been used in other contexts such as in reference to experimental manipulation of a meditation experience (e.g., Tan et al., 2014; Vinchurkar et al., 2014). Though no factor analysis was conducted in its development, the state-MAAS has shown sound psychometric properties, with an internal consistency of 0.92 and convergent validity with the trait-MAAS (Brown & Ryan, 2003).

The next most commonly used state mindfulness measure is the 13-item Toronto Mindfulness Scale (TMS; Lau et al., 2006). The TMS is theoretically grounded in a mix of sources, but mainly inspired by Bishop et al.'s (2004) description of mindfulness as attentional self-regulation and an orientation to experience the internal/external world with curiosity,

acceptance, and openness. The scale was found to be (through factor analysis) two-dimensional, consisting of (a) a “Curiosity” dimension reflecting one’s awareness of the present moment and whether that awareness is characterized by an open and curious stance (e.g., “I was curious about each of the thoughts and feelings that I was having”), and (b) a “Decentering” dimension reflecting being aware of one’s thoughts and feelings without being entangled in them (e.g., “I experienced myself as separate from my changing thoughts and feelings”). These items were designed for, and developed with, the experimental manipulation of a meditation experience (a 15-minute unguided meditation session in a community sample and amongst individuals with mindfulness meditation experience). Though the TMS has primarily been validated in populations with previous meditation experience, and in reference to an immediately preceding meditation session, the scale has since been used in other contexts such as among non-meditators, and in reference to various experimental tasks (e.g., watching a film clip; Erisman & Roemer, 2011). The factor structure of the TMS was assessed using exploratory (EFA) and confirmatory (CFA) factor analysis. It has sound psychometric properties, with an internal consistency of 0.86 and 0.87, for “Curiosity” and “Decentering,” respectively. As the development article of the TMS did not include testing whether the dimensions do or do not load onto a superordinate mindfulness construct, the creators imply keeping the two dimensions separate, rather than using a total score.

A less commonly used state mindfulness measure is the 21-item State Mindfulness Scale (SMS; Tanay & Bernstein, 2013). Like the TMS, the SMS is theoretically grounded in Bishop et al.’s (2004) definition of mindfulness, as well as in traditional Buddhist scholarship (specifically the Theravada Abhidhamma and the Satipatthana Sutta), and the items were developed with systematic feedback from mindfulness researchers and instructors. The scale was found to be (through factor analysis, see below) two-dimensional, consisting of (a) a “Mindfulness of Mind” dimension reflecting awareness of mental events including one’s thoughts and emotions (e.g., “I was aware of what was going on in my mind”), and (b) a “Mindfulness of Body” dimension reflecting awareness of one’s body sensations (e.g., “I clearly physically felt what was going on in my body”). The scale was developed using a single experimental manipulation (a mindfulness meditation versus a control task, in a student and community sample). The factor structure of the SMS was assessed using EFA and CFA. It has sound psychometric properties, with an internal consistency of 0.90 and 0.95, for “Mind” and “Body,” respectively. Because the two dimensions were found to load onto a superordinate

mindfulness construct, the SMS allows the use of either separate dimension scores or a total score.

The least commonly used state mindfulness measure is the 9-item Multidimensional State Mindfulness Questionnaire, which was developed in the German language (Blanke & Brose, 2017). Items for the MSMQ were chosen using a deductive approach, being drawn from the most commonly used mindfulness conceptions and measures based on reviews and citation counts; most of the tested items were adapted from a nonsystematic selection from the Cognitive and Affective Mindfulness Scale- Revised (CAMS-R; Feldman et al., 2007) and the trait-FFMQ (Baer et al., 2006) facets. Decisions about dimensionality were a mixture of the authors’ discretion (based on their reading of the previous literature) and results from their multilevel CFA, which led to the creation of a three-dimensional scale: (a) “Present-moment attention” (e.g., “I focused my attention on the present moment”), (b) “Acting with awareness” (e.g., “I sometimes did not stay focused on what was happening in the present”, reverse scored), and (c) “Nonjudgmental acceptance” (e.g., “Things went through my mind that I should not really be engaging myself with”, reverse scored). The scale was developed using an ESM design (70 students recorded six times per day for nine consecutive days). The MSMQ has adequate psychometric properties, with internal consistency ranging from 0.63 to 0.71 across the three dimensions. Though the three dimensions significantly load onto a superordinate mindfulness factor, the creators of the MSMQ do not comment on whether a total score calculation is recommended.

Finally, there have been sporadic attempts to create ad hoc state adaptations of the trait-FFMQ. This approach typically involves selecting a handful of items from certain trait-FFMQ facets and changing the wording to present tense and has been employed within various research designs (e.g., Eisenlohr-Moul et al., 2016; Frieze & Hofmann, 2016; Gavrilova & Zawadzki, 2023; Raugh et al., 2023; Snippe et al., 2015). Critically, however, none of these studies psychometrically validated their state adaptations.

Although each of the validated state mindfulness measures has merit, here we argue the need for a new state mindfulness measure, one that is specifically fashioned after the trait-FFMQ. We begin by describing what we believe are some of the limitations of the current state measures and then proceed to explain our rationale for moving forward with our new measure, referred to as the “state-4FMQ.” Broadly speaking, an ideal state mindfulness measure would (a) be brief (noting that the 21-item SMS is too long to be used practically in an ILD design), (b) use accessible language (noting that the TMS may not be relevant to meditation-naïve participants, and the

MSMQ has not been validated in English), (c) include an appropriate number, and type, of dimensions (noting that the state-MAAS has just one dimension, the relevance of the TMS dimensions have been questioned, and the MSMQ failed to conduct an EFA first to develop a empirically driven theory of dimensionality), (d) include both positive and reverse coded items within a dimension in an attempt to address response bias, improve construct representation, and increase generalizability (noting that no state mindfulness measure does this: the state-MAAS is entirely reverse coded, the TMS and SMS are entirely positively coded, and the MSMQ's Attention facet is entirely positively coded while its Awareness and Nonjudgement facets are entirely reverse coded), and (e) be versatile enough to accommodate references to both an in-lab manipulation as well as an ILD design. In creating a new state mindfulness measure, we aimed to accommodate all of these objectives and selected the trait-FFMQ as the basis to do so, based on the following reasons.

First, the trait-FFMQ is currently the most widely used multidimensional measure of mindfulness, and in a review of self-report mindfulness measures, it was regarded as providing the most comprehensive coverage of mindfulness for use in the general population (Bergomi et al., 2013). Second, past attempts to operationalize mindfulness that were *theoretically* derived produced exceedingly varied measurement scales. The development of the trait-FFMQ has a unique strength in that it was instead *empirically* derived via an EFA of 112 items collected from five independently developed self-report trait mindfulness measures, thereby consolidating a large pool of items from diverse sources through an unbiased statistical approach. In the same development article, the final EFA structure was then replicated in a 39-item hierarchical CFA showing a five-factor model with a superordinate mindfulness factor (i.e., the 39 items load onto five separate latent factors, and also those five factors themselves further load onto one superordinate latent factor). This data-driven finding of several (five) facets of mindfulness seems harmonious with the rich essence of mindfulness, as described in various Buddhist texts. Third, and related to the last point, the breadth of dimensions is critical since, as noted by Baer et al. (2004) in the development article of the Kentucky Inventory of Mindfulness Skills (a predecessor to the trait-FFMQ), different facets of mindfulness can uniquely predict a given dependent variable, and only a multidimensional measure can elucidate these differential relationships (e.g., see Medvedev et al., 2018; Petrocchi & Ottaviani, 2016; Prieto-Fidalgo et al., 2021 for articles that found selective predictive power of the different trait-FFMQ facets). By decomposing mindfulness into its constituent parts, researchers can better map out which facets of

mindfulness account for the most variance in the dependent variable in a given context. This can be particularly useful in an ILD design, where an additional decomposition of within versus between person effects for each facet can provide a more precise modeling of underlying mechanisms of action.

Thus, the present research aimed to provide initial psychometric validation of a new multidimensional state mindfulness measure adapted from the trait-FFMQ. Such a scale would be useful as it would allow researchers to better assess the mechanisms of state mindfulness on outcome measures and to tailor mindfulness interventions to target the specific facets most associated with increased well-being. The research was divided into two studies. In Study 1, we developed and explored the factor structure of the state items. Following in the footsteps of the development of the trait-FFMQ, we made no assumptions about the dimensionality of the underlying structure. In Study 2, we confirmed this factor structure in an independent sample and investigated the following additional forms of measurement validity for the state-4FMQ: convergent validity (testing whether the measure is related to established measures of the same construct), predictive validity (testing whether the measure is predictive of a dependent variable in an expected way), robustness (testing whether the predictive value of the measure remains robust after accounting for covariates that are related to the measure), and construct validity (testing whether a measure designed to assess a particular construct is actually measuring that construct in an expected way).

Study 1

Introduction

The aim of Study 1 was to create and provide initial validation and internal consistency metrics of a new state adaptation of the trait-FFMQ (the "state-4FMQ"). We hypothesized that multiple facets would emerge from the EFA results, but given inconsistencies in the trait-FFMQ's factor structure in the literature (e.g., Burzler & Tran, 2022), we had no *a priori* hypothesis about the specific number of facets or items that would emerge. We also hypothesized that the facets would themselves be moderately correlated and thereby suggestive of a superordinate mindfulness structure (tested in the CFA of Study 2, see below).

Method

We report how we determined our sample size, all data exclusion criteria, all manipulations, and all measures in the study.

Participants. Participants were undergraduate students recruited in 2022 through the UCSD participant pool, an online tool run by the Department of Psychology where undergraduate students sign-up to participate in research studies in exchange for course credit. Eligibility was restricted to participants who reported being at least 18 years old, able to complete the entire study in a private and quiet environment, having working audio on their computing device, and being comfortable listening to a 20-minute audio recording. The recruitment information, consent form, and protocol all referred to the study being about “relaxation” so as to not bias participants with the word “mediation” (see Dickenson et al., 2013). All participants gave their informed consent before participating and were compensated with course credit.

Combining the rule of thumb that sample sizes for EFA should be 5 to 10 participants per item, that our initial analysis involved 63 items, and that based on pilot data an estimated 15% of participants would be removed due to attrition and data cleaning, an initial sample of about 556 was needed. The collected sample consisted of 592 participants.

The following four exclusion criteria (as outlined in our pre-registration) were applied to the collected sample. *First*, 49 participants were excluded for failing to complete the entirety of the study (i.e., due to attrition). *Second*, 7 participants were excluded for failing to complete the study within ± 3 standard deviations of the median study duration. *Third*, 24 participants were excluded for failing to correctly respond to at least three out of four attention check questions dispersed throughout the survey. *Fourth*, 26 and 20 participants were excluded for admitting (at the end of the study) to not answering the survey questions honestly, or not fully engaging in the relaxation exercise, respectively (see item wording in Supplemental materials). In sum, a total of 126 participants were excluded for not passing these criteria. While we acknowledge that our exclusion criteria are strict and therefore limit the ecological validity of obtained results, we chose to prioritize data quality over generalizability. We felt this approach was necessary as the online nature of our study made it susceptible to participants not putting forth their best effort.

The final sample thus consisted of 466 participants between the ages of 18 and 45 years ($M = 20.66$, $SD = 2.91$). The majority reported being female at birth (78.76%) and reported their ethno-racial group as Asian (53.00%), followed by Hispanic or Latino (19.31%), White (13.30%), Mixed (7.30%), Middle Eastern or North African (4.51%), Black or African American (1.29%), Prefer not to say (0.89%), First Nation or Indigenous American (0.21%), and Native Hawaiian or Other Pacific Islander (0.21%).

Procedure. This study was conducted entirely online and remotely, and all data were collected via the survey program Qualtrics. All questions were required to be answered, so there were no missing values in the data. After filling out demographic questions including age, sex (assigned at birth), and ethno-racial identity, they were asked to ensure they had working audio and were in a private and quiet setting before proceeding with the “relaxation exercise,” consisting of listening to a 20-minute audio recording. Participants then started (by pressing a key) the audio recording, and they could not advance with the experiment until the audio was finished playing.

The audio recording consisted of a 20-minute mindfulness meditation recording, which guides participants to focus on the breath while accepting and letting go of thoughts and feelings as they arise, using instructions like “when you get distracted by either your thoughts or your feelings, simply notice the thought or feeling and return your focus back to the breath.” This meditation style is one of the most popular forms of meditation in the West and is commonly used for individuals without previous meditation experience. Note that the meditation script intentionally avoided using any phrases or keywords that were used in the subsequent state mindfulness questions to avoid contamination (see Supplemental Materials for a full transcription of the audio). As in our previous studies (Bondi, 2021), the meditation script was a joint effort between the second author and an outside professional, both of whom had years of experience guiding mindfulness meditations. The outside professional did the voicing in the recording.

At the end of the meditation, participants completed our new state mindfulness measure by reflecting on their experience during the immediately preceding “relaxation exercise.” Participants then answered a few extra demographic questions including items pertaining to previous meditation experience, and last, the “survey honesty” and “exercise engagement” questions (used as exclusion criteria; see above).

Measures

State-4FMQ. To create the state-4FMQ, we followed the procedure used by other researchers when creating state adaptations of trait measures (e.g., Neff et al., 2021). First, we rephrased all 39 items of the original trait-FFMQ to employ present moment language so that they would be relevant to any state situation, and generalizable enough to be used across study designs (e.g., an in-lab experiment, or an ILD). In addition, we included 24 items in the opposite direction (i.e., reversing some positively coded trait items and vice versa) to

ensure that each facet had the opportunity to have both positive and reverse-coded items. Though more research is needed regarding the effects of positive- versus reverse-phrased items (and other wording choices, particularly in the context of mindfulness measures, Karl & Fischer, 2020), including both is often recommended for reducing systematic response biases (e.g., Field, 2024; Weijters & Baumgartner, 2012, and see Grossman & Van Dam, 2011; Grossman, 2011 for arguments why using only reverse-phrased items, as in previous state mindfulness measures, might be problematic). Our research team piloted these items and further modified them to ensure they were accessible to individuals with or without previous meditation experience. This resulted in a total of 63¹ items to be tested in our study (see Supplemental Materials).

In the context of this study, the header text read, “Below are several questions about your experience during the 20-minute relaxation audio. Please respond with your honest opinion of what your experience was really like, rather than what you think your experience should have been.” All items were answered on a sliding scale from 1 to 5 with a resolution of 0.1 and with three labels: Not at all (1), Moderately (b), Completely/Entirely (c), with the order of items randomized to avoid potential order effects.

Meditation Status. This was used for descriptive purposes and included four questions. First, “Do you have any previous meditation experience (e.g., mindfulness meditation, transcendental meditation, loving-kindness meditation, etc.)?”, with Yes or No response options. Second, “About how frequently do you meditate?” with seven response options ranging from Several times a day, to Once a year or more. Third, “About how much time in minutes, on average, do you spend meditating per session?”, with nine response options ranging from 1–2 min to 60 + min. Fourth, “About how long have you been practicing meditation?”, with seven response options ranging from Less than 1 month, to 5 + years. In this study, we divided the sample into A) “current meditators” ($n = 71$; 15.25%) as participants that selected “Yes” for any previous experience, and once a week or greater for frequency of practicing; and B) “non-meditators” ($n = 395$; 84.76%) as all other participants. Note that the measurement of Meditation Status is examined more closely in Study 2.

Data Analysis. Basic descriptive analyses will report on means, standard deviations, and frequencies of demographics plus other relevant variables. To verify the univariate suitability of the data for EFA, the following metrics were considered for each individual item:

histograms revealing inappropriate distributions (e.g., obvious ceiling or floor effects, bimodal or truncated distributions); extreme skew ($>|1.0|$) or excess kurtosis ($>|1.5|$) values; and extreme bivariate correlational values ($r > |.80|$) with other items. To verify the multivariate suitability of the data for EFA, the following metrics with their commonly reported cutoffs were considered for the pool of items: the Kaiser-Meyer-Olkin (KMO) test with cutoffs of a minimum of .5; .5 to .7 are mediocre, .7 to .8 good, .8 to .9 great, .9 + superb; the Bartlett’s Test of Sphericity, where the p -value should be less than .05; and the Determinant of the correlation matrix, which should be greater than 0.00001. Determinant values lower than this minimum recommended threshold indicate multicollinearity or singularity in the data, which can be problematic in factor analysis (Field, 2024; Hair et al., 2019).

Since the items are continuous, the maximum likelihood estimate would have been used to run the EFA if the assumption of multivariate normality, as assessed with Mardia’s skewness and kurtosis tests (Mardia, 1970), was met. Because multivariate normality was not met (see Results), the more robust Weighted Least Squares estimate was used. An oblique (oblimin) rotation was used to permit co-varying multidimensional factors. Note that unlike the development article for the trait-FFMQ, all analyses were conducted on items as opposed to parcels. To determine the number of factors to extract, three different methods were used to ensure a more robust estimate: Kaiser-Guttman rule, Scree plot, and Parallel analysis.

Item removal strategies in EFA for multidimensional constructs, including cutoff criteria for loadings and cross-loadings, differ widely in the literature and lack a gold standard (Guvendir & Özkan, 2022). For this analysis, all of the above steps were iteratively repeated until an acceptable factorial solution was found. At each iteration, the first step was to eliminate all items that failed to load greater than or equal to 0.40 onto any one factor. If any items were eliminated due to this step, we ran the next iteration before checking the first step again. Only after all items passed this first step at the start of a new iteration would we consider a second step of eliminating items that showed a cross-loading of greater than or equal to 0.20, which would ensure that all retained items have a difference of at least 0.20 between the highest and next highest factor loadings. This second step was done one item at a time, from the highest to lowest cross-loading item. Given the large number of items in the initial pool reflecting overlapping concepts, content redundancy within a single factor was also considered during this process. To ensure that this item reduction process was empirically supported, we also reported on the Bayesian Information Criterion

(BIC) value of each iteration to ensure that each new iteration fits the data better than the previous iteration, which would be suggested if the BIC values approach 0 with each new iteration.

For the final iteration, we reported on model fit via the chi-square statistic (χ^2), Tucker-Lewis fit index (TLI), and root mean square error of approximation (RMSEA). However, we note these fit criteria are less strict and less relevant for EFA as compared to CFA analysis (Study 2). Factors were further evaluated for appropriateness including all factors having at least three items each; the conceptual interpretability of each factor; and the lack of suspiciously strong bivariate correlations between the factors. Furthermore, internal consistency for each factor and all items together was calculated with omega total and Cronbach's alpha statistics. While .70 or greater is a general rule of thumb for acceptable omega total and Cronbach's alpha statistics, we acknowledge that this cutoff should be expected to vary, for example, by the number of items per factor (Field et al., 2012).

All data were analyzed using R (Version 4.2.2; R Core Team, 2022).

Results

Preliminary Steps. As a preliminary step, all 63 potential items were examined for univariate suitability for an EFA. This was first assessed by detecting items where 25% or more of all participant responses were at the floor or ceiling of the response options (i.e., a 1 or a 5 on the 1–5 sliding scale). Two potential items adapted from Nonjudging were removed for being at the ceiling. No other items needed to be removed based on visual inspection of histograms to detect items with obviously non-normal distributions. Further, all of the remaining 61 items met acceptable ranges of skewness (range = -0.75 to 0.44), excess kurtosis (range = -1.25 to 0.30), and bivariate correlation values with other items (range = 0.10 to 0.75).

Next, we determined the multivariate suitability of the *pool* of 61 items for EFA. The Kaiser-Meyer-Olkin (KMO) value was 0.89 , suggesting that the sample size adequacy for running an EFA on the 61 items was great. Bartlett's Test of Sphericity was highly significant, $\chi^2(1830) = 13105.13$, $p < .001$, suggesting that the item pool was correlated enough to run an EFA. However, the determinant of the correlation matrix was 1.5×10^{-13} , suggesting the presence of a serious multicollinearity or singularity issue that was not detected with the bivariate item correlations. Note that this issue was unsurprising given that we purposely included a large number of interrelated items in the initial item pool as to ensure that all concepts embedded in the trait-

FFMQ were represented with a range of viable state-adapted options, and since an unexpectedly small number of items (2) were removed in tests of univariate suitability for EFA.

Because the aforementioned empirical tests of univariate suitability failed to eliminate most items and thus left too many viable options, we re-examined the pool of 61 items to further reduce the number of options. The goal was to arrive at a final set of items that (a) resulted in an acceptable determinant value, (b) retained comprehensive coverage of all content areas (including heterogeneous content areas within a factor) covered in the trait-FFMQ, and (c) included both positive and reverse coded items within each content area. In cases where there were several viable items with high content overlap, we favored items that: (a) were adapted from trait-FFMQ items that in a previous study had been shown to be highly sensitive to change across different contexts and therefore more reflective of state rather than trait tendencies (Truong et al., 2020), (b) had been selected for use in previously validated short-forms of the trait-FFMQ (Bohlmeijer et al., 2011; Gu et al., 2016; Tran et al., 2013), and (c) required minimal content modification when adapted from the trait-FFMQ, and (d) was determined to be of high theoretical relevance. This process, which we acknowledge has an inherent but unavoidable subjective bias, resulted in a reduced pool of 25 items that we believed accurately represented a state adaptation of the trait-FFMQ while simultaneously minimizing conceptual redundancy to a degree reflected in an acceptable determinant value (see Supplementary Materials for items). For this pool of 25 items, the KMO value was very good at 0.86 ; Bartlett's test of Sphericity was highly significant, $\chi^2(300) = 3847.669$, $p < .001$; and the determinant of the correlation matrix was 0.0002 , suggesting that EFA was now suitable to our dataset.

EFA Analysis. In all iterations, multivariate non-normality was indicated by Mardia's skewness and kurtosis tests, both being significant at $p < .001$, so the more robust weighted least squares (WLS) extraction method was used. In the first iteration, Kaiser rule, Scree plot, and Parallel analysis all suggested that five factors should be extracted. Thus, an EFA with an oblimin (oblique) rotation, using the weighted least squares extraction method, and five presupposed factors, was analyzed. The BIC of this iteration was -852.35 . At this iteration, we removed four items that failed to load greater than or equal to 0.40 onto any one factor: item 27, 29, 30, 47.

In the second iteration, Kaiser rule, Scree plot, and Parallel analysis all suggested that five factors should be extracted. Thus, an EFA with an oblimin (oblique)

Table 2. Exploratory Factor Solution of the state-4FMQ.

Item	Describing	ActAware	Nonjudging	Observing
1.	−0.03	0.82	0.00	−0.01
6.	−0.02	0.74	0.04	0.04
5.	0.11	0.71	−0.03	−0.06
2.	−0.05	0.60	0.06	0.16
13.	0.85	0.03	−0.03	−0.04
18.	0.78	−0.03	0.06	0.03
19.	0.78	0.08	−0.06	−0.10
15.	0.67	−0.08	0.04	0.13
17.	0.65	−0.03	0.06	0.14
35.	0.05	0.00	−0.03	0.70
33.	−0.02	0.02	0.07	0.66
38.	0.02	0.08	−0.02	0.60
36.	0.06	0.11	−0.14	0.40
52.	−0.03	−0.07	0.72	0.06
62.	0.02	0.04	0.71	0.02
59.	0.09	0.03	0.58	−0.11
55.	−0.08	0.00	0.56	−0.16
49.	0.07	0.15	0.56	0.07
α	.86	.82	.77	.71
ω	.89	.84	.79	.73
<i>Factor Correlation Coefficients</i>				
Describing	—	.24	.15	.31
ActAware		—	.23	.31
Nonjudging			—	.15
Observing				—

Note. α denotes Cronbach's alpha, ω denotes McDonald's omega. Factor loadings larger than or equal to 0.40 are in bold font. All items were positively keyed for analysis. As all state items aligned with their adapted factor from the trait-FFMQ, we used the trait-FFMQ's naming conventions for the retained factors.

rotation, using the weighted least squares extraction method, and five presupposed factors, was analyzed. The BIC of this iteration was −552.37. At this step, we removed two items which failed to load greater than or equal to 0.40 onto any one factor: item 21, 22. Note that this resulted in all state items derived from trait Nonreactivity being eliminated.

In the third iteration, Kaiser rule, Scree plot, and Parallel analysis all suggested that five factors should be extracted. Thus, an EFA with an oblimin (oblique) rotation, using the weighted least squares extraction method, and five presupposed factors, was analyzed. The BIC of this iteration was −432.27. At this step, we removed the one item that failed to load greater than or equal to 0.40 onto any one factor: item 8.

In the fourth and final iteration, Kaiser rule, Scree plot, and Parallel analysis all suggested that *four* factors should be extracted. Note that this was the first iteration that suggested four, rather than five, factors. Thus, an EFA with an oblimin (oblique) rotation, using the weighted least squares extraction method, and four presupposed factors, was analyzed. The BIC of this iteration was −343.49. All 18 items of this iteration were

retained since they each loaded greater than or equal to 0.40 onto any one factor, and since there were no substantial cross-loadings. Note that at each iteration the BIC value converged closer to zero, indicating a progressive improvement in data fitting throughout the iterative process.

Table 2 (upper panel) shows the standardized factor loadings of the four-factor model that emerged from these final 18 items. All loadings were at least moderately large, ranging from 0.40 to 0.85 in magnitude, indicating that the items converged meaningfully onto their respective factors. Given that the four-factor structure aligned entirely with the factors of the trait-FFMQ, we retained the original trait label to label each state factor that emerged from the EFA. In the final version, ActAware had 4 items; Describing had 5 items; Observing had 4 items; and Nonjudging had 5 items. The four-factor EFA had a good data fit, with the following indices all falling within acceptable cutoff criteria: $\chi^2(87) = 191.054$, $p < .001$; TLI = .934; RMSEA = .051, 95% confidence interval = [.041, .060].

Table 2 (middle panel) shows the internal consistency of the factors. All factors were identified by at least four

items and their internal consistency coefficients were satisfactory, ranging from .73 to .89 (MacDonald's ω). Furthermore, the internal consistency of all items together was high ($\omega = .87$).

Table 2 (lower panel) shows the factor correlations, clearly indicating nonzero relationships amongst all factors. This is consistent with the potential of a hierarchical solution (to be tested in the CFA in Study 2) and the plausibility of using a total score.

Study 1 Discussion

In Study 1 we found that a new 18-item, four factor state adaptation of the trait-FFMQ had a good data fit with satisfactory internal consistency. Notably, all state items aligned entirely with the adapted factors from the original trait-FFMQ, supporting the face validity of the state-4FMQ through a successful translation of trait to state items. Finally, correlations between factors were positive and significant as predicted, suggesting the possible existence of a factor structure with a superordinate dimension, to be tested in Study 2. The one unpredicted result was that all state items derived from trait Nonreactivity (and thus a state Nonreactivity factor) were eliminated during the EFA procedure, which we return to in the *General Discussion*.

Study 2

Introduction

Study 2 was conducted in four parts (described in detail, below), which allowed us to address two aims. First, we conducted a confirmatory factor analysis (CFA) to confirm the four-factor structure revealed in Study 1. For the CFA, we used all 18 items from the EFA (which we refer to as the *full-form state-4FMQ*) as well as a shortened version with 12-items (which we refer to as the *short-form state-4FMQ*, explained below). Second, data from Study 2 allowed us to further test the validity of the full- and short-form state-4FMQ, by testing for convergent, predictive, robustness, and construct validity.

Method

Participants. The participant pool and eligibility criteria were the same as described in Study 1, with the distinction that no individuals from Study 1 were included in Study 2. Unlike Study 1, all participants were tested in both a Meditation condition (as in Study 1) and a Control condition, randomized in order between Parts 1 and 3 (see Procedure, below). For the CFA, we only included data from participants that completed Parts 1-

2 (i.e., completion of Parts 3-4 was not required), and were tested with the Meditation first (in Part 1). Had we included data from participants who were tested with the Meditation second (in Part 3), we feared this could potentially contaminate the CFA results (as these participants would have already been familiarized with the state-4FMQ (or other mindfulness) items if tested in the Control condition first. This was of the utmost importance since this contamination could lead to inaccurate interpretations of the underlying factor structure, which was the basis for the further validity analyses. By contrast, for the "further validation analyses" (e.g., convergent, predictive, etc.), we included data from any participant that completed Parts 1-4. The reason for this is two-fold. First, for some validity analyses (i.e., construct validity) we had to use data from both Parts 1 and 3. Second, for our other validity analyses (e.g., predictive validity derived from just the Meditation condition), we were less concerned about contamination since they more broadly assess relationships between variables, rather than the specific structure of the measurement model. Therefore, while previous exposure to items is critical to avoid in CFA to maintain the integrity of the derived factor structure, its impact was less pronounced in our other validity tests. As such, we describe data collection, exclusion criteria, and demographics separately for these two segments (i.e., CFA vs. "further validation analyses").

First, for the CFA analysis, combining the rule of thumb that sample sizes should be 10–15 subjects per item in a confirmatory factor analysis (Pett et al., 2003), that our largest planned factor analysis involved 18 items, and that based on pilot data an estimated 15% of participants would be removed due to attrition and data cleaning, an initial sample of about 265 participants was needed. The collected sample consisted of 313 participants. The following five exclusion criteria were applied to the collected sample (as outlined in our pre-registration). *First*, 32 participants were excluded for failing to complete the entirety of Parts 1-2 (i.e. due to attrition) or failing to enter a matching participant ID between each of the two parts. *Second*, 9 participants were excluded for failing to complete the study within ± 3 standard deviations of the median study duration. *Third*, 6 participants were excluded for failing to correctly respond to at least two out of four attention check questions dispersed throughout the two parts. *Fourth*, 11 and 9 participants were excluded for admitting (at the end of the study) to not answering the survey questions honestly, or not fully engaging in the relaxation exercise, respectively. *Fifth*, 4 participants were excluded for being outliers, defined as scoring ± 3 standard deviations of the mean total state-4FMQ score. In sum, a total of 71 participants were excluded for not passing

these criteria. The final sample for the CFA analysis consisted of 242 participants between the ages of 18 and 46 years ($M = 21.45$, $SD = 3.69$). The majority reported being female at birth (85.54%) and most reported their ethnoracial group as Asian (50.00%), followed by Hispanic or Latino (24.79%), White (13.64%), Mixed (8.68%), Black or African American (2.07%), and Middle Eastern or North African (0.83%).

Second, for the further validity analyses, our sample size was based on an a priori power analysis for multiple linear regression calculated for the predictive validity analysis, with the following parameters: anticipated effect size $f^2 = .04$ (based on pilot data); statistical power = .80; four predictor variables in the predictive validity model; and an alpha of .05. This results in a sample size of 301 participants. With an estimated 15% of participants being excluded due to attrition and data cleaning, we therefore aimed to collect data from 354 participants. The collected sample consisted of 455 participants. The following four exclusion criteria (as outlined in our pre-registration) were applied to the collected sample. *First*, 75 participants were excluded for failing to complete the entirety of Parts 1–4 (i.e., due to attrition) or failing to enter a matching participant ID between each of the four parts. *Second*, 28 participants were excluded for failing to complete the study within ± 3 standard deviations of the median study duration. *Third*, 17 participants were excluded for failing to correctly respond to at least four out of eight attention check questions dispersed throughout the four parts. *Fourth*, 20 and 4 participants were excluded for admitting (at the end of the study) to not answering the survey questions honestly, or not fully engaging in either relaxation exercise, respectively. In sum, a total of 144 participants were excluded for not passing these criteria. The final further validity analyses sample thus consisted of 311 participants between the ages of 18 - 46 years ($M = 21.34$, $SD = 3.83$). The majority reported being female at birth (80.71%) and most reported their ethnoracial group as Asian (49.20%), followed by Hispanic or Latino (20.58%), White (17.36%), Mixed (8.68%), Black or African American (2.25%), Middle Eastern or North African (1.61%), and First Nation or Indigenous American (0.32%).

Procedure. This research was conducted entirely online and remotely, and all data were collected via the survey program Qualtrics. All questions were required to be answered, so there were no missing values in the data. The study consisted of four parts.

Part 1 (day 1): The order of events was as follows. First, participants filled out two state affect measures, which were randomized in order. Next, they were asked

to ensure they had working audio and were in a private and quiet setting before proceeding with a “relaxation exercise,” consisting of listening to a 20-minute audio recording. They then started (by pressing a key) the audio recording, and they could not advance with the experiment until the audio was finished playing. The audio recording was evenly randomized to consist either of the Meditation audio described in Study 1, or a Control audio consisting of an informative narration about the science and benefits of several relaxation techniques that has previously been used in our lab (Bondi, 2021) and was designed to be affectively neutral (see Supplementary Materials for a full transcription of the audio). Note that the script in the Control condition only mentioned relaxation techniques that could not be practiced in the moment, such as gardening and journaling. Though the two conditions (Meditation and Control) differed on structure, timing, and word count, they were matched for duration of instruction, beginning with the instruction of closing the eyes and relaxing, and being voiced by the same individual (which was the same person who narrated in Study 1). No order effects were observed; participants who received the Control condition first did not differ in demographic characteristics or predictive model outcomes from those who received the Meditation condition first. After the audio recording, they completed the following measures in this order: the state-4FMQ, the two other state mindfulness surveys in randomized order, the two state affect measures in randomized order, the clarification of instructions question, and last, the survey honesty and exercise engagement questions.

Part 2 (day 2): One day later, the order of events was as follows. First, participants completed measures of trait mindfulness and chronic stress in randomized order, then they answered standard questions about demographics, and last, the survey honesty question.

Part 3 (day 8): One week after Part 1, participants repeated the procedure from Part 1 but were assigned to listen to the audio recording that they did not listen to in Part 1 (e.g., if they were assigned to listen to the Meditation audio in Part 1, they were assigned to listen to the Control audio for Part 3).

Part 4 (day 9): One day after Part 3, participants repeated the procedure from Part 2, but were asked to report on previous meditation experience instead of standard demographics.

Measures

State Mindfulness. The following state mindfulness measures were obtained following the relaxation exercises and instructed participants to reflect on their experience during the immediately preceding “relaxation exercise” (see above).

State-4FMQ. The 18-item State Four Facet Mindfulness Questionnaire from Study 1 was our primary measure of state mindfulness. The items were presented randomly. The scale showed acceptable internal consistency (see Results).

State-MAAS. The State Mindful Attention and Awareness Scale (state-MAAS; Brown & Ryan, 2003, reviewed in the *General Introduction*), a five-item scale designed to measure a recent or current expression of the mindful attention and awareness of one's engagement in daily activities, is currently the most commonly used measure of state mindfulness. All items are reverse coded and rated on a 7-point Likert scale with three labels: Not at all (0), Somewhat (3), and Very Much (6). Items are averaged to calculate a total score. Mindfulness is conceptualized as a unidimensional construct. The scale showed acceptable internal consistency with Cronbach's alpha coefficients ranging from $\alpha = .84$ to $.86$ in the present study

SMS. The State Mindfulness Scale (SMS; Tanay & Bernstein, 2013, reviewed in the *General Introduction*), a 21-item scale comprising of a mindfulness of mind subscale, a mindfulness of body subscale, and a total score. Items are rated on a 5-point Likert scale from 1 = Not at all to 5 = Very Much. Subscales and the total score are averaged. The scale showed acceptable internal consistency with Cronbach's alpha coefficients ranging from $\alpha = .90$ to $.93$ for the total score, and $\alpha = .79$ to $.90$ for the subscales, in the present study.

State Affect. The following state affect measures were obtained before and after the relaxation exercises and instructed participants to reflect on their experience "right now, in this moment".

State Anxiety. The state items from the State-Trait Anxiety Inventory (STAI; Spielberger, 1983). This 20-item test measures the presence and severity of current symptoms of anxiety. It is set up as a 4-point Likert-scale from 1 ("Not at All") to 4 ("Very Much So"). State Anxiety was used because relief from anxiety is one of the most widely promoted benefits of mindfulness (e.g., Russ et al., 2017; Van Dam et al., 2018). The scale showed acceptable internal consistency with Cronbach's alpha coefficients ranging from $\alpha = .93$ to $.94$ in the present study.

State Stress. As no validated State Stress measure could be found in the literature, a composite score was calculated by combining the responses to three in-house questions. The first two questions have slider scales from 1 to 7 with a resolution of 0.1, with labels on each end

and number markers in between, and the third question uses a 5-point visual analog scale. The first question asks: "How stressed do you feel right now?" (1 = not at all stressed, 7 = extremely stressed). The second question similarly asks: "How relaxed do you feel right now?" (1 = not at all relaxed, 7 = extremely relaxed) and is reverse coded. The third question asks: "How are you feeling right now?" and has a simple, traditional, yellow-and-black smiley face that can be adjusted from the neutral middle (starting point) to up to 2 points in the positive direction (making the face slightly, and then fully, smile) or up to 2 points in the negative direction (making the face slightly, and then fully, frown). State Stress was then calculated with the following equation: $(Q1 + (8-Q2) + (1.4*(6-Q3)))/21$, such that a higher score indicates more stress. Even though the third item is not continuous, the composite score is calculated with a resolution of 0.1. State Stress was used because mindfulness-based techniques are one of the most used coping strategies to handle stress (e.g., Aguilar-Raab et al., 2021; Weinstein et al., 2009). These three in-house items showed acceptable internal consistency with Cronbach's alpha coefficients ranging from $\alpha = .82$ to $.86$ in the present study, which supports the use of the composite score as a measure of State Stress.

Other State Measures

Thought Valence. At the end of the relaxation exercises, participants were asked the following: "Imagine someone read the transcript of what you thought about in the 20-minute exercise. How would they rate the content of that transcript?", which is answered on a 7-point Likert scale and includes 7 markers ranging from very negative to very positive (coded as 0 to 6). The wording of this in-house item was inspired by a previous ESM study (Gross et al., 2024) that found significant associations between Thought Valence, mindful attention, and mood. As a sanity check for this in-house item, participants were also asked to rate their *confidence* in their Thought Valence rating (i.e., "How confident are you about your estimate above?", from Not at all confident (0) to Extremely confident (6)), and the mean ratings were relatively high (Meditation $M = 4.14$, $SD = 0.96$; Control $M = 4.02$, $SD = 0.94$).

Clarity of Instructions. Since guided meditations may seem esoteric or inaccessible to novice meditators, we include an additional question asked in Feldman et al. (2010) as a manipulation check to ensure the instructions were clear to participants after the relaxation exercises: "To what extent did you feel that the audio recording instructions were clear enough for you to understand what you were being asked to do?", rated on

a 7 point Likert scale with three labels: 1 = Not at all, 4 = Somewhat, 7 = To a great extent.

Trait Measures. *Trait Mindfulness:* The 39-item trait-Five Facet Mindfulness Questionnaire (Baer et al., 2006) was used to measure trait mindfulness. Note that we use the 39-item version, and not an abbreviated version, since more knowledge about its psychometric properties is available in the literature, and since it is the basis for the state-4FMQ. The scale showed acceptable internal consistency with Cronbach's alpha coefficients ranging from $\alpha = .90$ to $.94$ for the total in the present study.

Chronic Stress: The Perceived Stress Scale (PSS; Cohen et al., 1983) is a 10-item scale of stress felt in the past month. The scale showed acceptable internal consistency with Cronbach's alpha coefficients ranging from $\alpha = .86$ to $.87$ in the present study.

Meditation Status. This is an important construct to consider in any study involving meditation since the dramatic increase in popularity of meditation and related practices, such as yoga or smartphone led mindfulness practices, means that much of the population has at least some familiarity with meditation (see Burzler et al., 2019; Heppner & Shirk, 2018). Studies that rely on student or community samples (as opposed to targeted recruitment strategies such as reaching out to monks or MBSR instructors) usually rely on participant self-report measures of previous meditation experience that often include items on the frequency, duration, and/or type of meditation practiced. However, these studies use vastly heterogeneous classification approaches with no current consensus as to best practices (for a discussion on issues arising from different definitions of meditation experience, see Davidson & Kaszniak, 2015; Van Dam et al., 2018). To be consistent with a commonly used classification in the literature (e.g., Baer et al., 2008; Burzler et al., 2019; Feldman et al., 2010; Schlosser et al., 2022), we operationalized "current meditators" being participants that reported currently practicing meditation or mindfulness at least once a week. Informed by Pang and Ruch (2019), we further operationalized "past meditators" as those who practiced at least once a week but no longer do so. All others were categorized as "non-meditators" (see Supplemental Materials for all item descriptions). Note that for simplicity we refer simply to meditators, rather than those with meditation or mindfulness experience. We acknowledge that stricter criteria involving how long one has been practicing for, the duration of each practice, and the type of practice, could be used in future studies involving samples from targeted populations.

Data Analysis. Basic descriptive analyses reported on means, standard deviations, and frequencies of relevant variables. Normality, as assessed with visual inspection of histograms, was verified and met for all variables of interest. The assumptions of all statistical tests were checked and met. The level of significance was set to 5% ($p < .05$) for all tests; however, we emphasized effect sizes rather than statistical significance since the latter is often misleading. Effect sizes were reported as the following: Pearson r values for bivariate correlations, with the rule of thumb that absolute values of $.10$ to $.30$ are weak effects, $.30$ to $.50$ are medium effects, and $.50$ and over are large effects (Cohen, 1988, p. 198); Cohen's d for t -tests, with the rule of thumb that values around $.20$ are considered small effects, values around $.50$ are considered medium effects, and values around $.80$ or more are considered large effects (Cohen, 1988); Cramer's V for chi-square tests, with the rule of thumb that values $\geq .1$ are weak, $\geq .3$ are moderate, and $\geq .5$ are large effects (Kakudji et al., 2020); and partial eta squared (η^2) for analysis of variance (ANOVA) and regression models, with the rule of thumb that $\eta^2 = .01$ indicates a small effect; $\eta^2 = .06$ indicates a medium effect; and $\eta^2 = .14$ indicates a large effect (Cohen, 1988). All ANOVA and regression models in this study use Type III sum of squares, which examines individual effects in light of all other model effects regardless of order. All data were analyzed using R (Version 4.2.2; R Core Team, 2022).

CFA Analysis. At 18-items, the state-4FMQ from Study 1 was relatively lengthy compared to most state mindfulness measures. We ultimately wanted the state-4FMQ to be as brief (yet comprehensive) as possible to facilitate its use in research designs with multiple measures and/or measures administered on multiple occasions, without overburdening participants. Inspired by previous literature and on the results of the EFA from Study 1, we thus created the short-form state-4FMQ using the same sample as the full-form (as has been done in the development of other measures, for example, Ullrich-French et al., 2021). The selection of which items to retain involved empirical (e.g., items with the highest factor loadings), conceptual (i.e., items that when grouped together minimized redundancy and maximized conceptual coverage within a factor), and theoretical (e.g., referencing the extant literature on the short-forms of the trait-FFMQ) considerations. We aimed to retain three items per factor, which is generally considered the minimum number of items needed for model identification (Kline, 2015). We hypothesized that the 12-item short-form would retain the psychometric properties of, and produce comparable results with, the full 18-item state-4FMQ. To

test this, all of the following analyses were conducted on both the full (18-item) and short (12-item) state-4FMQ.

The main goal of the CFA was to confirm the accuracy of the four-factor structure of the state-4FMQ that emerged from Study 1 through EFA. Prior to conducting the CFA, we first calculated bivariate correlations amongst the state-4FMQ facets. Based on the results of Study 1 and the expectation that the facets would be different enough to be separable facets (as would be revealed in the four-factor CFA models), yet similar enough to be associated via one superordinate mindfulness construct (as would be supported by a hierarchical CFA structure), weak to moderate positive correlations between all facets were predicted. We also assessed the internal consistency of the state-4FMQ with Cronbach's alpha (α) and McDonald's omega (ω) statistics for each facet. Note that while 0.70 or greater is a general rule of thumb for acceptable estimates for both statistics, we point out that this cutoff can be expected to vary, specifically, being lower when the number of items per factor is small (Field et al., 2012). We also point out that there is some controversy in the appropriate range of acceptable alpha values (Streiner, 2003), and the argument has been made that lower cutoff values are appropriate in early stages of scale development (Nunnally, 1967). For these reasons, we believe that promoting the brevity of a measure is a fair tradeoff for relatively reduced internal consistency values.

To verify the suitability of the data for CFA, the state-4FMQ items were checked for univariate normality via extreme values for skewness ($>|1.0|$) and excess kurtosis ($>|1.5|$) and visual inspection of histograms. The items were also checked for multivariate normality with Mardia's multivariate skew and kurtosis tests (Mardia, 1970). Using the R-package lavaan (Rosseel, 2012), we employed a CFA on these items to test a four-factor solution. Replicating the development of the trait-FFMQ (Baer et al., 2006), error terms were not allowed to covary and items were constrained to load onto only one factor in accordance with the theorized measurement model. Unlike the development of the trait-FFMQ, CFA models used individual items as opposed to item parceling, since the latter is a controversial and less stringent practice and is not advisable when there are small numbers of items per factor (as in the current study). Further, one study found that using an item-level CFA of the trait-FFMQ produced comparable results to parceling (Christopher et al., 2012), suggesting there is little added value to the parceling method.

We hypothesized that (a) the results of the CFA would confirm a multidimensional structure with four distinct factors (i.e., a four-factor solution without a superordinate mindfulness factor, which we refer to as

the "four-factor nonhierarchical" model), and (b) in a four-factor hierarchical CFA, the factors would strongly load onto a superordinate mindfulness factor (i.e., a four-factor solution with a superordinate mindfulness factor, which we refer to as a "hierarchical" model). The finding of a hierarchical structure would suggest that the four facets are sufficiently interrelated to be considered part of one overarching mindfulness construct. We hypothesized, however, that Observing might not strongly load onto the superordinate factor because our population was likely to have minimal to no meditation experience, and Baer et al. (2006) among others found that trait-Observing only significantly loaded onto a superordinate factor amongst individuals with sufficient meditation experience. The following indices, with suggested benchmarks by Hu and Bentler (1999) and Brown (2015), were evaluated collectively to provide an evaluation of how well each model fit the data: chi-square statistic (χ^2 ; $p > .05$), root mean square error of approximation (RMSEA; <0.06) with its 90% confidence interval (0.00–0.08), standardized root mean square residual (SRMR; <0.08), Tucker-Lewis fit index (TLI; >0.90), and comparative fit index (CFI; >0.95).

Besides assessing individual model fit, we also aimed to determine the overall best fitting factor structure. In addition to our two main models described above, we were inspired by the moderate bivariate correlations of the facets we observed in the EFA, as well as factor structures that have been tested with the trait-FFMQ, to also compare the fit of the following additional models: a one-factor model containing only a total mindfulness score (which assumes that the state-4FMQ has a unidimensional structure), and an exploratory four-factor bifactor model (which is similar to the four-factor hierarchical model with the exception that latent variables are set as orthogonal to each other, and all items simultaneously load on one general factor and four specific factors). We hypothesized that the one-factor model would provide the worst fit, but had no *a priori* hypothesis about which of the three other models (i.e., the four-factor nonhierarchical, hierarchical, or bifactor model) would provide the best fit given the inconsistent factorial structures found for the trait-FFMQ (see Burzler & Tran, 2022). Note that determining the best-fitting factor structure of the state-4FMQ affects the scoring guidelines of the measure. Specifically, a four-factor hierarchical or bifactor model would support the use of both facet scores and a total score. By contrast, a four-factor nonhierarchical model would support the use of individual facets but not a total score, and a one-factor model would support only the use of a total score. Model comparisons mainly utilized Bayesian Information Criterion (BIC) values, though we will also reported on the Akaike information criterion (AIC). For BIC and AIC values,

scores closer to 0 indicate a more parsimonious and better-fitting model. Though relevant differences between BIC values are rules of thumb, we used the following guidelines proposed by Raftery (1995): differences >10 as very strong evidence, 6 to 10 as strong evidence; 2 to 6 as positive evidence; and 0 to 2 as weak evidence, for a model being a better fit.

Further Validation Analyses. All analyses were conducted on both the full (18-item)- and short (12-item)-form state-4FMQ. Because the ultimate goal of this research was to create a viable *short* form measure of state mindfulness, in the Results, we present only results for the short-form (results using the full-form state-4FMQ were comparable to the short-form and are available in Supplemental Materials). Also note that some tests of construct validity (discriminant sensitivity and test-retest reliability) required using data from both the Meditation and Control conditions. By contrast, the convergent and predictive validity analyses included data from only the Meditation condition.

Convergent Validity. We assessed convergent validity via bivariate correlations between the state-4FMQ and several other measures. First were the two extant measures of state mindfulness, which were predicted to show weak to moderate positive correlations with the state-4FMQ facets, although the ActAware component of the state-4FMQ was predicted to be more strongly correlated with the state-MAAS since the latter focuses mainly on awareness. Though not preregistered, we later hypothesized a strong correlation between the state-4FMQ's Observing subscale and the SMS, in line with previous results measuring this association with the trait-FFMQ's Observing subscale (e.g., Navarrete et al., 2023; Tanay & Bernstein, 2013).

Second was the trait-FFMQ, where the aligned facets, for example, state ActAware and trait ActAware, should be more positively correlated than non-aligned facets, for example, state ActAware and trait Nonjudging. The correlations between the state- and trait-FFMQ should not be too high, however, else one could argue that our new state measure is behaving in a trait-like way. There is good theoretical reasons (Robinson & Clore, 2002a, 2002b) and empirical data showing rather weak associations between trait mindfulness and induced mindful states (Bravo et al., 2018; Heppner & Shirk, 2018) to suggest that our analysis should likewise show only moderate correlations.

Predictive Validity. Here, we tested whether any (or all) of the facets of the state-4FMQ predict state affect, in the form of changes in State Anxiety and State Stress.

This was tested separately for each of the two state measures, since Cronbach's alpha statistics suggested they ought not be combined into a single metric (see Results). Using multiple linear regression, the post-intervention state affect score was the dependent variable, the pre-intervention state affect score was a covariate, and the state-4FMQ facets were entered simultaneously as predictor variables. Because our state affect measures are "negative" states, we expected a negative relationship between the state-4FMQ and state affect. Note that, in addition to providing a test of predictive validity, these multiple linear regressions allowed us to ask an empirical question—which state mindfulness facet best predicts state affect, in particular, state stress and anxiety, which we address by comparing effect size confidence intervals between facets.

It is perhaps important to explain our decision to obtain both a pre- and post-intervention score for state affect (with the pre-score being used as a covariate in the regression model) yet only a post-intervention score for the state-4FMQ. First, we wanted to reduce the threat to internal validity that can occur when assessing the same items in close proximity (Campbell & Stanley, 1963) and felt this was particularly important for the state-4FMQ items, as they are the items being used to develop our new scale. Second, had we attempted to obtain pre-intervention scores for the state-4FMQ items, it would have been difficult to standardize the pre-intervention reference point (e.g., responding in reference to what activity the participant was doing just before the study began), which is critical since previous research has shown that the degree of mindfulness varies across different daily activities (e.g., Gross et al., 2024; Killingsworth & Gilbert, 2010; Raynes & Dobkins, 2025). This was not an issue for the dependent measure (i.e., the state affect scores), for which participants were instructed to report on how they were feeling "right now" (both pre- and post-intervention).

We hypothesized that ActAware and Nonjudging would most strongly predict state affect, followed by Describing, and that Observing would not be predictive. This prediction was guided by two recent meta-analyses reporting on correlational studies of the trait-FFMQ (Carpenter et al., 2019; Mattes, 2019). Of course, our results may be expected to differ from these reviews since we measure state, rather than trait, mindfulness; and since we use an experimental design to measure the benefit of a single meditation, rather than a correlational design to measure general associations between mindfulness and external referents. We also note the current research differs from most extant experimental designs investigating the benefits of meditation since most of these studies conduct interventions over the course of days, weeks, or even months at a time. Studies exploring

the short-term effects of a single meditation session are a more sparse but growing area of research that generally suggests that there are likely measurable benefits to one session of meditation, albeit with relatively small and transient effects (e.g., see reviews by Gill et al., 2020, p. 202; Schumer et al., 2018; Williams et al., 2023).

Robustness. Given that we found evidence for predictive validity (see Results), in the next step, we tested whether the strength of the relationship between state-4FMQ and state affect remained robust when two other covariates were included in the model, namely the trait-FFMQ and Thought Valence. In particular, checking the strength of the relationship between the state-4FMQ and state affect in the presence of the trait-FFMQ is one important step in demonstrating the construct validity of the state-4FMQ. Before conducting this robustness analysis, we first tested whether either covariate (trait-FFMQ total score and Thought Valence) had the potential to account for some (or all) of the predictive validity results by looking at bivariate relationships between the state-4FMQ facet scores and the two covariates, and between the two covariates and state affect. We further confirmed whether a potential covariate interacted with the state-4FMQ facet scores in predicting state affect; if it did interact, that term was removed from the model.

Construct Validity. Here, we asked if the state-4FMQ was behaving in a way that was consistent with its intended design of (a) being a “State” versus Trait” measure, as well as (b) detecting “mindful” states. Evidence that the state-4FMQ was behaving in a state-like way was addressed in two ways. First, when testing for robustness of the state-4FMQ (above), we reasoned that if the relationship between the state-4FMQ and state affect remained strong while accounting for the effects of the trait-FFMQ, this would demonstrate that the new state-4FMQ is not masquerading as a trait measure. Second, data from the two times points over which the state-4FMQ was collected (i.e., after the Meditation and Control conditions, i.e., parts 1 and 3, tested 1 week apart), allowed us to assess its test-retest reliability. If the state-4FMQ is truly a state measure, then its test-retest reliability should be substantially lower in comparison to that of the trait-FFMQ (the latter obtained from two time points over which the trait-FFMQ was collected, i.e., parts 2 and 4, tested 1 week apart). To be comprehensive for these analyses, we conducted bivariate correlations of the scores for the following measures (which were each taken at two timepoints, one week apart): two trait measures (PSS, trait-FFMQ), and three state mindfulness measures (state-4FMQ, SMS, and state-MAAS). Note that we calculate bivariate Pearson correlations, rather than repeated measures correlations,

since the one-week time gap between paired administrations was reasonably long enough to meet the assumption of independence of observations.

Second, to investigate whether the state-4FMQ detects *mindful* states, we conducted a test of “discriminant sensitivity” wherein state-4FMQ scores were expected to be higher after the Meditation versus the Control condition. This was implemented using a mixed ANOVA with the state-4FMQ Total score as the dependent variable, and Condition (measured within-subject) entered as the predictor variable. In addition, we included Meditation Status (measured between-subjects) and its interaction with condition as terms in the model, since previous studies have shown that Meditation Status in a student sample can impact responses to meditation exercises (e.g., Thompson & Waltz, 2007), or scores on measures related to mindfulness (e.g., Baer et al., 2008); specifically, we were concerned that discriminant sensitivity to the Meditation condition might only be detected amongst current and past meditators, but not non-meditators.

The study design, hypotheses, and analysis plan were prospectively pre-registered (prior to data collection) and are publicly available here: <https://osf.io/xm49a>.

Results

CFA Analysis. Descriptive statistics and bivariate correlations for the state-4FMQ facets are presented in Table 3. The means of all facet scores were approximately at the midpoint (3.00) of the scale for both the full-form (range $M = 2.99$ – 3.74) and short-form (range $M = 2.85$ – 3.75) state-4FMQ.

In the full-form state-4FMQ, we found weak to moderate correlations (r 's(240) = .17–.41, p 's < .01) between most facets, with the exception of one insignificant relationship between Observing and Nonjudging (r (240) = .07). This suggests that the facets represent related but distinct constructs and supports the possibility of a hierarchical CFA solution. The short-form state-4FMQ demonstrated a similar pattern of results, with one additional insignificant relationship between ActAware and Nonjudging (r (240) = .09).

Internal consistency reliability scores indicated that all facets and the total of the state-4FMQ demonstrated adequate to good internal consistency with the exception of Observing in the short-form state-4FMQ, which still approached an acceptable value ($\alpha = .66$; $\omega = .67$). As expected given its abbreviated length, the short form of each variable had marginally lower internal consistency than its corresponding full form. Thus, the internal consistency of both the full and short state-4FMQ was determined to be satisfactory.

Table 3. Descriptive Statistics, Internal Consistency, and Bivariate Correlations of the State-4FMQ.

Facet	M (SD)	α	Ω	1	2	3	4
1. ActAware	2.99 (0.87)/2.85 (0.91)	.85/.83	.87/.84				
2. Describing	3.17 (0.84) /3.19 (0.89)	.89/.83	.91/.83	.35**/.30**			
3. Nonjudging	3.55 (0.71) /3.61 (0.84)	.71/.70	.76/.72	.17**/.09	.29**/.34**		
4. Observing	3.74 (0.62) /3.75 (0.64)	.73/.66	.76/.67	.39**/.39**	.41**/.38**	.07/.01	
5. Total	13.45 (2.09) /13.40 (2.18)	.85/.79	.89/.86	.73**/.98**	.77**/.78**	.55**/.56**	.65**/.61**

Note. M represents mean, SD represents standard deviation, α represents Cronbach's alpha, and ω represents McDonald's omega. Results for the full-form state-4FMQ are presented left of the dash, and for the short-form state-4FMQ on the right of the dash. All correlations had 240° of freedom.

* $p < .05$. ** $p < .01$.

Table 4. Model Fit and Comparison Indices.

Model	Model Fit Indices					Model Comparison Indices	
	RMSEA [90% CI]	SRMR	TLI	CFI	χ^2 (df)	BIC	AIC
<u>Full-form state-4FMQ</u>							
One Factor	0.15 [.14–.16]	0.13	0.49	0.55	745.74 (135)	11524.82	11399.22
Four-Factor Nonhierarchical	0.06 [.05–.08]	0.07	0.91	0.93	231.49 (129)	10929.12	10782.58
Four-Factor Hierarchical	0.07 [.05–.08]	0.07	0.91	0.92	239.71 (131)	10930.60	10791.05
Four-Factor Bifactor	0.06 [.05–.08]	0.07	0.91	0.93	209.82 (117)	10973.28	10784.88
<u>Short-form state-4FMQ</u>							
One Factor	0.19 [.17–.21]	0.15	0.41	0.51	436.51 (54)	7783.95	7700.21
Four-Factor Nonhierarchical	0.06 [.04–.08]	0.05	0.94	0.95	88.14 (48)	7386.81	7282.14
Four-Factor Hierarchical	0.08 [.06–.09]	0.08	0.91	0.93	109.03 (50)	7402.76	7305.07
Four-Factor Bifactor	0.08 [.06–.10]	0.08	0.89	0.93	101.77 (42)	7439.51	7313.91

Note. RMSEA = Root Mean Square Error of Approximation; SRMR = Standardized Root Mean Square Residual; TLI = Tucker-Lewis fit index; CFI = Comparative Fit Index; χ^2 = Scaled Chi-Square Statistic; df = degrees of freedom; BIC = Bayesian Information Criterion; AIC = Akaike information criterion.

All χ^2 statistics were significant at $p < .001$.

Preliminary analyses showed that all 18 items of the state-4FMQ were approximately normally distributed, as assessed by levels of skewness (range -0.71 to 0.23) and excess kurtosis (range -1.12 to 0.39), and visual inspection of histograms. A violation of multivariate normality was indicated by Mardia's skewness and kurtosis tests both being significant at $p < .001$ for both the full and short state-4FMQ. Since this multivariate normality assumption was not met in our sample, rather than conducting CFAs using the default maximum likelihood estimation, we instead used a maximum likelihood estimation with robust standard errors and a Satorra-Bentler scaled test statistic, which is less dependent on the assumption of normality (Li, 2016).

For both the full and short state-4FMQ, the four-factor nonhierarchical, hierarchical, and bifactor models demonstrated acceptable fit indices (RMSEA $< .08$; CFI/TLI $> .90$, see Table 4) based on Hu and Bentler (1999), even though they just missed our strictest a priori cutoffs. The one-factor model failed to approach any acceptable model fit indices with the exception of the scaled chi-square statistic, suggesting that this model was a poor representation of the data, and was therefore not considered in the model comparisons below.

In terms of model comparisons, which was addressed with BIC values, for both the full and short state-4FMQ, the four-factor nonhierarchical model fit better than the four-factor hierarchical model (full Δ BIC = 1.48 ; short

Table 5. Confirmatory Factor Analysis of the Full and Short Four-Factor Nonhierarchical Models.

Factor	Item	Loading
ActAware	1. My mind was wandering off and I was easily distracted	0.79 (0.83)
	5. It was easy to pay attention and focus on what I was doing	0.75 (0.72)
	6. I found it difficult to stay focused on what was happening in the moment	0.84 (0.81)
	2. I didn't pay attention to what I was doing because I was daydreaming, worrying, or otherwise distracted	(0.72)
Describing	13. I would have been good at finding the words to describe my feelings	0.73 (0.80)
	15. It would have been hard for me to find the words to describe what I was thinking	0.85 (0.81)
	17. Finding the right words to describe the sensations in my body would have been difficult for me	0.78 (0.78)
	18. I would have been able to find a way to put my feelings into words	(0.78)
	19. It would have felt natural to put my experience into words	(0.73)
Observing	33. I failed to notice sensations in my body	0.71 (0.81)
	36. I noticed how the experience affected my thoughts, bodily sensations, and emotions	0.53 (0.47)
	38. I paid attention to sensations	0.66 (0.61)
	35. I did not have much awareness of bodily sensations	(0.65)
Nonjudging	52. Some of my thoughts were abnormal or bad and I shouldn't have thought in that way	0.89 (0.81)
	59. My emotions were normal and there was no need to change the way I was feeling	0.51 (0.55)
	62. I disapproved of myself for having irrational ideas	0.59 (0.64)
	55. I did not consider whether my thoughts were good or bad	(0.37)
	49. Regardless of my thoughts or emotions, I accepted myself	(0.49)

Note. Loadings are standardized and are all significant at $p < .001$. Items 1, 2, 6, 15, 17, 33, 35, 52, and 62 should be reversed before scoring. All items were positively keyed for analysis. Items retained in the short form are bolded. Loadings for the full-form state-4FMQ are in parentheses.

$\Delta\text{BIC} = 15.95$), and the four-factor hierarchical model fit better than the four-factor bifactor model (full $\Delta\text{BIC} = 42.86$; short $\Delta\text{BIC} = 36.75$). Though differences were more obvious in the short than the full state-4FMQ, the four-factor nonhierarchical model demonstrated marginally more favorable results for all other fit and comparison indices as well. Taken together, the pattern of findings suggests that although the four-factor hierarchical model fit reasonably well, the four-factor nonhierarchical model provided the optimal fit for both the full ($\chi^2 (129) = 231.49$, $p < .001$, $\text{RMSEA} = 0.06$ (90% CI 0.05–0.08), $\text{SRMR} = 0.07$, $\text{TLI} = 0.91$, $\text{CFI} = 0.93$) and short-form ($\chi^2 (48) = 88.14$, $p < .001$, $\text{RMSEA} = 0.06$ (90% CI 0.04–0.08), $\text{SRMR} = 0.05$, $\text{TLI} = 0.94$, $\text{CFI} = 0.95$) state-4FMQ.

When comparing the short versus full state-4FMQ results, the short-form state-4FMQ clearly provided more favorable fit and comparison indices across all models. However, this was expected since the short-form state-4FMQ is nested within the full-form state-4FMQ, and therefore by default has fewer parameters to estimate and is less prone to overfitting.

As the four-factor nonhierarchical model of the state-4FMQ was the best-fitting model, Table 5 displays summary statistics for this model only.² For the short-form state-4FMQ, all item loadings were both statistically significant and at least moderately large in magnitude on their respective factor (0.51–0.89, p 's $< .001$). Results

were similar for the full-form state-4FMQ, with the exception of one relatively weaker loading from item 55 (0.37; $p < .001$). Though this loading is not seriously problematic, it does fall beneath the expected 0.40 loading value for items in a well-fitting CFA model and lends credence to utilizing the short (which omits this item), rather than full, state-4FMQ.

Though we do not display the results of the four-factor hierarchical model given its inferiority to the four-factor nonhierarchical model, we report that in the four-factor hierarchical model each factor, including Observing, loaded significantly (p 's $< .01$) and at least moderately in magnitude on the superordinate factor for the full (0.41 to 0.77) and short (0.42–0.93) state-4FMQ.

Further Validation Analyses

Convergent Validity. Convergent validity was tested by investigating whether the state-4FMQ (a) shares variance with other state mindfulness measures and (b) shares variance with the trait-FFMQ, more so for aligned vs. misaligned facets. Bivariate correlations of all state and trait mindfulness³ measures are presented in Table 6. As predicted, there was a strong correlation between the state-MAAS and the ActAware facet of the state-4FMQ ($r(309) = .69$), as well as some smaller correlations between the state-MAAS and the other three facets. The SMS subscales were somewhat strongly

Table 6. Convergent Validity of the Short-form state-4FMQ.

Variable	1	2	3	4	5	6	7	8	9	10	11	12
1. State ActAware												
2. State Describing	.33**											
3. State Nonjudging	.15**	.37**										
4. State Observing	.39**	.36**	.10									
5. State-MAAS	.69**	.39**	.30**	.44**								
6. SMS Body	.26**	.31**	.08	.60**	.33**							
7. SMS Mind	.35**	.34**	.03	.62**	.34**	.67**						
8. Trait ActAware	.39**	.31**	.26**	.26**	.47**	.20**	.19**					
9. Trait Describing	.16**	.61**	.27**	.23**	.21**	.23**	.29**	.34**				
10. Trait Observing	.22**	.23**	.14*	.43**	.25**	.47**	.39**	.12*	.32**			
11. Trait Nonjudging	.18**	.17**	.51**	.09	.25**	-.04	.06	.42**	.28**	.04		
12. Trait Nonreactivity	.31**	.14*	.23**	.17**	.26**	.11	.18**	.28**	.23**	.36**	.38**	

Note. Correlations hypothesized to be moderate to strongly correlated are bolded. All correlations had 309 degrees of freedom. There were no duplicate items between the state-4FMQ and any additional measure. To simplify this table, the following variables were omitted but are available by request from the authors: PSS, SMS total, Trait-FFMQ Total, state-4FMQ Total.

* $p < .05$. ** $p < .01$.

correlated with state Observing ($r's(309) = .60-.62$). Last, there existed moderate correlations between the aligned facets of the state and trait-FFMQ ($r's(309) = .39-.61$), which were strongest for aligned vs. misaligned facets. (Also note that, as was expected based on the results from the CFA analysis, the facets of the state-4FMQ were weakly to moderately correlated with one another ($r's(309) = .15-.39$), with the exception of one insignificant relationship between Observing and Nonjudging ($r(309) = .10$).

Predictive Validity. Predictive validity was tested by asking if any (or all) facets of the state-4FMQ predicted two other state measures, specifically, State Stress and State Anxiety (both of which can be thought of as “negative” affective states). Given the conceptual overlap of State Anxiety and State Stress (plus the fact that our State Stress measure was “in-house”), we first tested whether the two were correlated enough to be combined into a single measure. Although the correlation of the two measures was strong ($r(309) = .82$, $p < .001$), the Cronbach’s alpha of all the State Anxiety and State Stress items together was not suspiciously high ($\alpha = .68$) as to assume completely overlapping constructs. Thus, we kept the two measures separate for the remainder of our multiple linear regression models. In these models, the dependent variable was State Stress (or State Anxiety) measured *post*-Meditation and the predictor variables were those same metrics measured *pre*-Meditation plus the four facets of the state-4FMQ. The results are summarized in Table 7 and Table 8 (left panels), noting that the two different dependent variables yielded similar results.

As would be expected, pre-Meditation scores were strong predictors of post-intervention scores ($p's < .001$, $\eta^2 = .23$ to $.35$), which itself provides evidence for the reliability of the measures. With respect to state-4FMQ facets, all but one (negatively) predicted State Anxiety/State Stress. Specifically, ActAware and Nonjudging were significant predictors with medium effect sizes ($p's < .001$, $\eta^2 = .04$ to $.05$), followed closely by Observing with slightly smaller effect sizes ($p's = .002-.001$, $\eta^2 = .03-.04$), while Describing was not significant in either model ($p's = .19-.27$).

Robustness. Here, we asked whether the predictive validity of the facets seen in the above analysis remained robust when trait-FFMQ and Thought Valence were added to the models. Both of these variables were found to be suitable as covariates as they correlated with both the main predictor variables (i.e., the state-4FMQ facets) and the dependent variables (i.e., state affect). Specifically, the trait-FFMQ Total score positively correlated with all state-4FMQ facet scores ($r's(309) = .35$ to $.46$, $p's < .001$) and negatively correlated with both dependent variables ($r's(309) = -.40$ to $-.35$, $p's < .001$). Similarly, Thought Valence positively correlated with all state-4FMQ facet scores ($r's(309) = .18-.32$, $p's < .001$), indicating that thoughts were more positive for participants who experienced higher levels of state mindfulness, and negatively correlated with both dependent variables ($r's(309) = -.34$ to $-.33$, $p's < .001$). Before proceeding with our robustness analyses, we also ensured that neither covariate interacted with any state-4FMQ facet in the models.

As shown in Tables 7 and 8 (right panels), there were no notable differences in the beta coefficient, p -value, or

Table 7. Multiple Linear Regression Results Predicting State Anxiety After a Single Meditation Session.

Variable	Predictive Model				Robustness Model			
	<i>B</i>	<i>SE</i>	<i>P</i>	η^2	<i>B</i>	<i>SE</i>	<i>p</i>	η^2
(Intercept)	40.65	2.62	<.001		40.22	3.08	<.001	
Pre-Score	0.36	0.03	<.001	.35	0.37	0.03	<.001	.35
				[.27, .42]				[.27, .42]
ActAware	−1.22	0.33	<.001	.04	−0.99	0.33	.003	.03
				[.01, .10]				[.00, .07]
Describing	−0.47	0.36	.19	.006	−0.54	0.36	.13	.007
				[.00, .03]				[.00, .04]
Nonjudging	−1.49	0.37	<.001	.05	−1.47	0.37	<.001	.05
				[.01, .11]				[.01, .10]
Observing	−1.44	0.45	.002	.03	−1.37	0.45	.002	.03
				[.00, .08]				[.00, .08]
Trait-FFMQ Total					0.02	0.02	.20	.005
								[.00, .03]
Thought Valence					−0.91	0.23	<.001	.05
								[.01, .10]
Adj. R^2				0.55				0.58

Note. $N = 311$. The F-statistic is significant at $p < .001$ in all models. *B* represents unstandardized regression weights. η^2 represents partial eta-squared. Numbers in brackets represent 95% confidence intervals. Collinearity as assessed with VIF was well within acceptable limits for each model. Predictors were entered into analysis simultaneously. Significant effects are bolded.

Table 8. Multiple Linear Regression Results Predicting State Stress After a Single Meditation Session.

Variable	Predictive Model				Robustness Model			
	<i>B</i>	<i>SE</i>	<i>P</i>	η^2	<i>B</i>	<i>SE</i>	<i>p</i>	η^2
(Intercept)	9.54	0.86	<.001		10.11	1.00	<.001	
Pre-Score	0.37	0.04	<.001	.23	0.37	0.04	<.001	.23
				[.15, .31]				[.16, .31]
ActAware	−0.54	0.13	<.001	.05	−0.42	0.13	.002	.03
				[.01, .11]				[.00, .08]
Describing	−0.16	0.14	.27	.004	−0.14	0.14	.34	.003
				[.00, .03]				[.00, .03]
Nonjudging	−0.51	0.15	<.001	.04	−0.45	0.15	.002	.03
				[.01, .09]				[.00, .08]
Observing	−0.67	0.18	<.001	.04	−0.61	0.18	.001	.04
				[.01, .10]				[.01, .09]
Trait-FFMQ Total					−0.001	0.01	.92	.00004
								[.00, .01]
Thought Valence					−0.34	0.09	<.001	.04
								[.01, .10]
Adj. R^2				0.43				0.46

Note. $N = 311$. The F-statistic is significant at $p < .001$ in all models. *B* represents unstandardized regression weights. η^2 represents partial eta-squared. Numbers in brackets represent 95% confidence intervals. Collinearity as assessed with VIF was well within acceptable limits for each model. Predictors were entered into analysis simultaneously. Significant effects are bolded.

effect size in any state-4FMQ facet after accounting for trait-FFMQ and Thought Valence, indicating that the predictive validity of the state-4FMQ is robust. The results of this analysis also revealed that while Thought Valence showed significant effects, that is, more positive thoughts were associated with less anxiety and stress (p 's < .001, $\eta^2 = .04-.05$), the trait-FFMQ Total did

not (p 's = .20–.92), the latter being relevant to construct validity, addressed in the next section.

Construct Validity

State versus Trait Aspects. The robustness analysis, above, provided one source of evidence that the state-

Table 9. Test-Retest Reliability of State and Trait Measures.

Measure	Correlation Between Timepoints
Chronic Stress	.83
Trait-FFMQ Total	.85
Trait ActAware	.82
Trait Describing	.83
Trait Nonjudging	.81
Trait Observing	.77
Trait Nonreactivity	.77
State-4FMQ Total	.62
State ActAware	.50
State Describing	.58
State Nonjudging	.53
State Observing	.41
State MAAS	.54
SMS Total	.50
SMS Mind	.47
SMS Body	.48

Note. All correlations were significant at $p < .001$ and had 309 degrees of freedom.

4FMQ behaves in a state- versus trait-like fashion. This is because the relationship between state-4FMQ and state affect remained strong while accounting for the effects of the trait-FFMQ, and moreover, the trait-FFMQ did not significantly predict state affect in this model despite their moderate bivariate association.

The second way we addressed the state-like aspect of the state-4FMQ was to compare test-retest reliability between the state-4FMQ and the trait-FFMQ. As shown in Table 9, the stability of the state-4FMQ facets across the two timepoints was substantially lower (r 's(309) = .41–.58, p 's < .001) than the trait-FFMQ facets (r 's(309) = .77–.83, p 's < .001), and was on par with the correlations of the state-MAAS and SMS.

Discriminant Sensitivity to Mindful States. Here, we tested construct validity of the state-4FMQ by asking if it detects mindful states. If so, state-4FMQ scores were expected to be higher following the Meditation condition (designed to induce a mindful state) as compared to the Control condition (which ought not to induce a mindful state). Because this analysis included Meditation Status as a moderator, we categorized our sample into three groups: current meditators ($n = 43$, 13.83%), past meditators ($n = 43$, 13.83%), and non-meditators ($n = 225$, 72.35%), noting the large differences in sample size were expected given our convenience sample of undergraduate students (for descriptive purposes, summarized responses to all meditation experience questions are described in Table 10 for current and past meditators).

The results from a mixed-ANOVA, in which the dependent variable was total state-4FMQ score, and the

predictor terms were Condition (Meditation vs. Control), Meditation Status (3 levels, see above) and the interaction of the two, we found no significant main effect of Condition (Meditation $M = 3.37$, $SD = 0.56$; Control $M = 3.45$, $SD = 0.59$), $F(1, 614) = 0.03$, $p = .86$, $\eta^2 < .001$), nor an interaction between Condition and Meditation Status ($F(2, 614) = 0.21$, $p = .81$, $\eta^2 = .001$). There was, however, a small but significant main effect of Meditation Status, $F(2, 614) = 3.80$, $p = .02$, $\eta^2 = .01$). Post-hoc pairwise comparisons using the Tukey method revealed that Current meditators exhibited a small but significantly higher total score than non-meditators (mean difference = 0.19, $SE = 0.07$, $p = .02$), and trended towards higher total scores than past meditators (mean difference = 0.17, $SE = 0.09$, $p = .12$). No significant difference was found between non-meditators and past meditators ($p = .98$). Exploratory analyses revealed a similar pattern of results was using each state-4FMQ facet score as the dependent variable.

In sum, we did not find evidence for discriminant sensitivity of the state-4FMQ, even when we considered Meditation Status, a factor that could have moderated this construct (noting that Meditation Status itself had a significant main effect in an expected direction). Because of the null finding, we conducted several extra exploratory analyses to rule out other potential explanations, which are presented in Supplemental Materials. As we return to the *Discussion*, we believe the null finding likely resulted from the Control condition being too similar to the Meditation condition (i.e., wherein the Control condition inadvertently induced a mindful state).

Study 2 Discussion

The results of the CFA confirmed the results of the EFA. Bivariate correlations across the state-4FMQ facets were mostly as expected; internal consistency scores were acceptable; a four-factor nonhierarchical CFA model displayed acceptable fit indices and was comparatively better fitting than other model specifications; and a short-form of the state-4FMQ produced comparable results as the full-form. The results of the model comparisons best support the use of individual facet scores but not a total score. However, because the four-factor hierarchical model fit reasonably well, and all factors including Observing significantly loaded onto the superordinate factor despite most of our sample being non-meditators, the use of a total score should not yet be ruled out.

In further validation analyses, the state-4FMQ was found to demonstrate convergent and predictive validity, and robustness. Results of construct validity were varied; whereas we found substantial evidence that the

Table 10. Descriptive Results of Meditation Experience Items for Current and Past Meditators.

	Current (N = 43)	Past (N = 43)	Overall (N = 86)
Frequency of practice			
Several times a day	10 (23.3%)	0 (0%)	10 (11.6%)
Once a day	5 (11.6%)	8 (18.6%)	13 (15.1%)
Several times a week	16 (37.2%)	16 (37.2%)	32 (37.2%)
Once a week	12 (27.9%)	19 (44.2%)	31 (36.0%)
Time per session (minutes)			
Mean (SD)	15.7 (9.86)	16.9 (11.6)	16.3 (10.7)
Length practicing (months)			
Mean (SD)	14.3 (10.6)	6.16 (5.91)	10.3 (9.49)
Type^a of practice (Mean %, SD %)			
Focused Attention Meditation	33.7 (34.4)	56.4 (34.7)	45.1 (36.2)
Loving-Kindness or Compassion Meditation	12.8 (19.8)	6.51 (16.4)	9.67 (18.4)
Open Monitoring Meditation	6.19 (13.2)	5.47 (13.8)	5.83 (13.4)
Mantra or Transcendental Meditation	1.53 (4.15)	3.95 (13.1)	2.74 (9.72)
Yoga, Tai Chi, or Qi Gong	20.3 (29.8)	16.7 (28.1)	18.5 (28.8)
Meditation-based Religious Practices	15.1 (27.8)	3.26 (13.0)	9.16 (22.4)
Not Sure	7.33 (23.8)	4.77 (15.2)	6.05 (19.9)
Other	3.02 (15.5)	2.91 (15.4)	2.97 (15.4)

^aParticipants assigned a percentage to each type of practice they engage(d) in, could select multiple options, and selections had to sum to 100%. Therefore each cell reflects the mean percentage assigned to each type across all participants in that column, and each column sums to 100%. Overall, most participants (70.93%) selected more than one type of meditation ($M = 2.42$, $SD = 1.26$).

state-4FMQ acts more like a state than a trait, we could not yet confirm its discriminant sensitivity to detecting mindful states.

General Discussion

In two studies employing exploratory factor analysis (EFA), confirmatory factor analysis (CFA), and several additional validation analyses, the current research provides initial psychometric support for the State Four Facet Mindfulness Questionnaire (state-4FMQ), a novel measure of state mindfulness adapted from the Five Facet Mindfulness Questionnaire (trait-FFMQ; Baer et al., 2006). This new measure addresses some missing aspects of the currently employed state mindfulness scales. Specifically, the state-4FMQ is short (12 items for the short-form), multi-dimensional (4 facets), contains reversed- (as well as positive-) scored items within each facet, created to be applicable across study designs (e.g., an in-lab experimental manipulation, or an ILD) and robust in a group of participants who are predominantly non-meditators (roughly 75% of our samples self-reported as not currently meditating). Overall, most of our hypotheses were confirmed, and the results support a theoretically based and psychometrically valid factor structure with four distinct facets closely aligned with those of the trait-FFMQ. For the remainder of the discussion, we address two notable unexpected results; the absence of a Nonreactivity factor in the EFA, and a null finding in the analysis of discriminant sensitivity.

This is followed by a summary of our empirical findings that are unrelated to scale development per se, specifically, determining which state mindfulness facets best predict state affect. We then end with a discussion of the limitations of the current research and make suggestions for future directions.

The Omission of Non-Reactivity. Although EFA is an atheoretical analysis and we therefore had no a priori hypotheses about the state-4FMQ's emergent factor structure, it was nonetheless surprising that all state items derived from trait Nonreactivity, and therefore a fifth Nonreactivity factor, were eliminated in the EFA. There are, however, several findings from the literature that support this omission. First, a previous attempt to create a state Nonreactivity factor as part of the MSMQ development article (Blanke & Brose, 2017) also eliminated Nonreactivity from the final model. The authors noted that Nonreactivity items are the most difficult to adapt from their trait origin given most trait items refer to specific distressing experiences unlikely to occur during a normal day, thereby necessitating substantial content modification. Second, attempts to replicate the trait-FFMQ's factor structure have sometimes demonstrated problems with its Nonreactivity factors, such as one study using a general population sample showing that a four-factor solution omitting Nonreactivity fit the data best (Solem et al., 2015), as well as studies showing more general psychometric issues including low internal

consistency of the factor or relatively weak item loadings onto the factor (e.g., Baer et al., 2008; Tran et al., 2013). As such, Tran et al. (2013) called for seriously revising or removing Nonreactivity items from the trait-FFMQ. Third, there has been conceptual criticism of including Nonreactivity as a core component of mindfulness in the general population. Nonreactivity has been described as a construct that only results from continued mindfulness practice (e.g., Baer et al., 2012); as an advanced skill to which all other mindfulness components are necessary preconditions (Burzler & Tran, 2022); as something better understood as a cognitive skill (Tran et al., 2013) or an outcome of mindfulness practice (Bishop et al., 2004) and as being population dependent, such that populations without sufficient meditation or mindfulness experience lack conceptual understanding of item content (e.g., Lecuona et al., 2020). Therefore, while a state Nonreactivity facet may have been retained had we used a targeted sample of experience meditators, its omission in our predominantly non-meditating sample strengthens the generalizability of our measure since it ensures greater relevance for future research using diverse population groups. Last, as there is some evidence that self-reported impulsivity (which can be seen as inversely related to Nonreactivity) is not effectively captured outside of ecological contexts (see Strickland & Johnson, 2021), the same may apply to Nonreactivity.

Null Results for Discriminant Sensitivity. The other unexpected result was a null finding in an analysis of the state-4FMQ's discriminant sensitivity to mindful states, which we believe is likely due to a limitation in the design of the Control condition used to test this. Specifically, we were very careful to match the Control condition (which described the scientific benefits of different relaxation techniques) in as many ways as possible to the Meditation condition, with the only difference being that only the latter was designed to induce a mindful state (see Method). In doing so, we had hoped to employ a rigorous methodological approach that minimizes confounding variables and enhances the ability to measure effects specific to Meditation. However, it is possible that our desire to match the two Conditions so closely may have resulted in them being too similar, and thus, their effects largely indistinguishable from one another (see Supplementary Materials for further information).

Specifically, the Control condition might have induced heightened mindfulness if participants were *imagining* themselves benefiting from the various relaxation techniques described in the audio, including listening to music, exercising, gardening, and journaling, any or all of which might be considered mindful activities.

The audio also explicitly encouraged them (at the end, for two minutes) to imagine “other” ways of reducing stress, which may have resulted in them choosing to meditate or imagine another mindful activity. Unfortunately, the current study lacked additional checks to ensure that our conditions were inducing the intended subjective experience to participants (i.e., the Control being relaxing and engaging, but not mindful; and the Meditation being mindful). Future studies can explore the utility of the Meditative State Scale (López et al., 2022), the Early Meditation Hindrances scale (Russ et al., 2017), open-ended qualitative items, or semi-structured interviews to better validate whether participants experienced each condition as intended by researchers. Furthermore, the discriminant sensitivity of the state-4FMQ to mindful states should be tested in studies designed to detect larger differences between conditions, such as a multi-day intervention using multiple comparison groups including a passive control of listening to an excerpt from an intentionally dull book (e.g., Zeidan et al., 2015).

It is also important to acknowledge that we did not apply bootstrapping techniques to test the moderation effect of the mixed ANOVA used in this test of discriminant sensitivity. Although bootstrapping is not typically employed in mixed ANOVAs, a growing body of literature supports its use. For example, although Preacher and Hayes (2008) focused on mediation, their arguments advocating for bootstrapping as a robust alternative to traditional statistical approaches are applicable to moderation analyses. Future research should thus consider incorporating bootstrapping methods to further validate these findings.

Differential Effects of Mindfulness Dimensions on State Affect. Although the main goal of the current study was to develop a new state mindfulness scale, the results from the predictive validity analysis allowed us to ask an empirical question; are certain dimensions of state mindfulness, more than others, associated with state affect? To date, this question has been investigated with trait mindfulness (e.g., Carpenter et al., 2019; Mattes, 2019), and is now only starting to be investigated with state effect (see Ullrich-French et al., 2021). In the current study, we found unique contributions of ActAware and Nonjudging, followed closely by Observing, to state affect, in particular, state stress and anxiety, after a single meditation session. Interestingly, the Describing facet did not uniquely contribute, which may have been due to the meditation style we used not focusing on the “labeling” of affective experiences (see Tran et al., 2013). Note that in another study in our lab (Raynes & Dobkins, 2025), we asked this question in an ILD design

and found similar results. More broadly, the relative contribution of each facet likely varies based not only on the experimental context (e.g., mediation type), but also on the study population (e.g., in terms of previous meditation experience), and the specific dependent variable used. Because affect is a multifaceted construct, its improvement encompasses not only the amelioration of negative states like stress or anxiety but also the enhancement of positive states such as happiness or contentment, future inquiries can build upon our findings by expanding the measurement of state affect to encompass both positive and negative dimensions. Such findings have implications for designing mindfulness trainings and interventions by suggesting which facets should be preferentially targeted for improvement of different aspects of wellbeing (see a similar point in Mattes, 2019).

Limitations. On a final note, we acknowledge some general limitations. As with any scale development article, creating the state-4FMQ necessarily involved some subjective decision-making. We developed items by adapting those from the trait-FFMQ (so they would fit any situational context), and in doing so, could have altered the construct underlying those items. Though we tried to avoid semantic overlap in the retained items, several items do repeat certain words, which may have artificially inflated factor loadings and internal consistency coefficients. However, this is a common issue in self-report measures, and we believe the current redundancy in phrasing is still substantially less than in the trait-FFMQ. When narrowing the pool of potential state-4FMQ items in the EFA, there was a tradeoff between trying to maximize factor loadings and internal consistency coefficients (e.g., by only selecting items with the strongest bivariate correlations) versus optimizing validity (e.g., by retaining a mixture of positive and reverse-coded items within a dimension). That said, we welcome authors to refer to our initial 63-item pool in Supplementary Materials and explore alternative approaches to developing a state-4FMQ. Moreover, given inconsistencies in factor structures commonly reported for the trait-FFMQ even within the same population type (see a review by Lecuona et al., 2020), more advanced statistical techniques could have been employed, such as exploratory structural equation modeling to model a hierarchical structure with two superordinate factors (as described for the trait-FFMQ in Burzler et al., 2019; Tran et al., 2013, 2014); as well as the inclusion of modification indices to suggest post hoc model modifications (e.g., allowing error terms to covary). Last, the state-4FMQ was developed using traditional psychometric methods (e.g., factor analysis and

validity testing). In addition to replicating the current results, future research could apply alternative methods including Item Response Theory, Differential Item Functioning Analysis, Rasch Modeling, and Generalizability Theory; noting that some of these methods require assessments across several timepoints, which was not collected in the current research. Using both traditional and newer psychometric methods will allow researchers to leverage the strengths of each approach and thus enhance the overall quality of the state-4FMQ as a self-report measurement instrument.

Another limitation of the current research involves the choice of sample population. Our data were obtained from undergraduate students who were disproportionately represented in terms of ethnicity (which can also be seen as a strength given the predominantly non-White composition of the sample), sex at birth (mostly female), and Meditation Status (mostly non-meditators), which impacts the generalizability of our findings. Therefore, the state-4FMQ should be tested with more diverse populations which can then be leveraged across various analytical methods. Specifically, since prior research has demonstrated that the structure and correlates of mindfulness components can vary across different population types (e.g., Baer et al., 2006; Bravo et al., 2018; Tran et al., 2013), future research should apply configural invariance testing on the four-factor nonhierarchical structure across a range of dimensions (e.g., Meditation Status, sex, education) as well as evaluate the stability of the factor structure within a sample over time such as before and after participation in a mindfulness-based intervention program. Furthermore, qualitative methods could be used to better understand how individuals with more disparate backgrounds (both demographics and of Meditation Status) interpret items. Meditation Status is a key dimension to further scrutinize in this regard since knowledge about mindfulness may influence how one interprets mindfulness self-report items (e.g., Grossman & Van Dam, 2011). However, this construct has been inconsistently operationalized in the literature with no consensus as to best practices; even with our relatively homogeneous sample, the descriptive results of all meditation experience questions revealed substantial variability across participants.

Despite these limitations, the results from the current studies provide initial support for a valid and reliable measure of state mindfulness inspired by the trait-FFMQ, which is composed of four distinct facets. While work remains to be done in validating and refining this measure, the state-4FMQ is worthy of further investigation as it has the potential to be a valuable measurement tool for researchers operationalizing mindfulness as a multidimensional state construct.

Acknowledgments

Not Applicable

Declaration of Conflicting Interests

The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Funding

The author(s) received no financial support for the research, authorship, and/or publication of this article.

Ethical Considerations

The Institutional Review Board at UCSD reviewed and approved the study protocol prior to commencement (approval #201994).

Consent to Participate

Written informed consent was obtained electronically from all individual participants included in the study.

Consent for Publication

Not applicable

Data Availability

All data are available at the Open Science Framework (<https://osf.io/xm49a>).

ORCID iD

Stephen Raynes  <https://orcid.org/0009-0000-7709-1236>

Supplemental Material

Supplemental material for this article is available online.

Notes

1. We intentionally started with a conservatively large pool of potential items since we expected many items to be eliminated early on in the EFA analysis.
2. Summary statistics of all models are available by request from the authors
3. Note that while we used the trait scores from Part 2, the results were similar if using the trait scores from Part 4.

References

- Aguilar-Raab, C., Stoffel, M., Hernández, C., Rahn, S., Moessner, M., Steinhilber, B., & Ditzen, B. (2021). Effects of a mindfulness-based intervention on mindfulness, stress, salivary alpha-amylase and cortisol in everyday life. *Psychophysiology*, 0(0), e13937. <https://doi.org/10.1111/psyp.13937>
- Baer, R. A., Carmody, J., & Hunsinger, M. (2012). Weekly Change in Mindfulness and Perceived Stress in a Mindfulness-Based Stress Reduction Program: Weekly Change in Mindfulness and Stress in MBSR. *Journal of Clinical Psychology*, 68(7), 755–765. <https://doi.org/10.1002/jclp.21865>
- 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. <https://doi.org/10.1177/1073191105283504>
- Baer, R. A., Smith, G. T., Lykins, E., Button, D., Krietemeyer, J., Sauer, S., Walsh, E., Duggan, D., & Williams, J. M. G. (2008). Construct Validity of the Five Facet Mindfulness Questionnaire in Meditating and Nonmeditating Samples. *Assessment*, 15(3), 329–342. <https://doi.org/10.1177/1073191107313003>
- Baer, R., Smith, G., & Allen, K. (2004). Assessment of Mindfulness by Self-Report: The Kentucky Inventory of Mindfulness Skills. *Assessment*, 11, 191–206. <https://doi.org/10.1177/1073191104268029>
- 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. <https://doi.org/10.1007/s12671-012-0110-9>
- Bishop, S. R., Lau, M., Shapiro, S., Carlson, L., Anderson, N. D., Carmody, J., Segal, Z. V., Abbey, S., Specia, M., Velting, D., & Devins, G. (2004). Mindfulness: A proposed operational definition. *Clinical Psychology: Science and Practice*, 11(3). <https://doi.org/10.1093/clipsy.bph077>
- Blanke, E., & Brose, A. (2017). Mindfulness in Daily Life: A Multidimensional Approach. *Mindfulness*, 8, 737–750. <https://doi.org/10.1007/s12671-016-0651-4>
- Bohlmeijer, E., Ten Klooster, P. M., Fledderus, M., Veehof, M., & Baer, R. (2011). Psychometric properties of the five facet mindfulness questionnaire in depressed adults and development of a short form. *Assessment*, 18(3), 308–320.
- Bondi, T. (2021). *Mechanisms of meditation: investigating the components and covariates of a single session of meditation* (Unpublished doctoral thesis, University of California San Diego, advisor: Karen Dobkins)
- Bravo, A. J., Pearson, M. R., Wilson, A. D., & Witkiewitz, K. (2018). When traits match states: examining the associations between self-report trait and state mindfulness following a state mindfulness induction. *Mindfulness*, 9(1), 199–211. <https://doi.org/10.1007/s12671-017-0763-5>
- 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. <https://doi.org/10.1037/0022-3514.84.4.822>
- Brown, K. W., & Ryan, R. M. (2004). Perils and promise in defining and measuring mindfulness: Observations from experience. *Clinical Psychology: Science and Practice*, 11(3), 242–248. <https://doi.org/10.1093/clipsy.bph078>
- Brown, K. W., Ryan, R. M., & Creswell, J. D. (2007). Mindfulness: theoretical foundations and evidence for its salutary effects. *Psychological Inquiry*, 18(4), 211–237. <https://doi.org/10.1080/10478400701598298>
- Brown, T. A. (2015). *Confirmatory factor analysis for applied research* (2nd ed., pp. xvii, 462). The Guilford Press.

- Buhk, A. H., Schultz, H., & Bullock, W. A. (2023). What are they measuring? Testing the convergent validity of state and trait mindfulness measures across two independent samples. *Journal of Psychopathology and Behavioral Assessment*, 45(4), 1107–1118. <https://doi.org/10.1007/s10862-023-10091-3>
- Burzler, M. A., & Tran, U. S. (2022). Dispositional mindfulness and the process of mindfulness cultivation: A qualitative synthesis and critical assessment of the extant literature on the five facet mindfulness questionnaire (FFMQ). *Collabra: Psychology*, 8(1), 56176. <https://doi.org/10.1525/collabra.56176>
- Burzler, M. A., Voracek, M., Hos, M., & Tran, U. S. (2019). Mechanisms of mindfulness in the general population. *Mindfulness*, 10(3), 469–480. <https://doi.org/10.1007/s12671-018-0988-y>
- Campbell, D. T., & Stanley, J. C. (1963). *Experimental and quasi-experimental designs for research*. Wadsworth.
- Carpenter, J. K., Conroy, K., Gomez, A. F., Curren, L. C., & Hofmann, S. G. (2019). The relationship between trait mindfulness and affective symptoms: A meta-analysis of the Five Facet Mindfulness Questionnaire (FFMQ). *Clinical Psychology Review*, 74, 101785. <https://doi.org/10.1016/j.cpr.2019.101785>
- Christopher, M. S., Neuser, N. J., Michael, P. G., & Baitman-galkar, A. (2012). Exploring the psychometric properties of the Five Facet Mindfulness Questionnaire. *Mindfulness*, 3(2), 124–131. <https://doi.org/10.1007/s12671-011-0086-x>
- Chu, S. T.-W., & Mak, W. (2020). How mindfulness enhances meaning in life: A meta-analysis of correlational studies and randomized controlled trials. *Mindfulness*, 11, 177–193. <https://doi.org/10.1007/s12671-019-01258-9>
- Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed.). Routledge. <https://www.taylorfrancis.com/books/mono/10.4324/9780203771587/statistical-power-analysis-behavioral-sciences-jacob-cohen>
- Cohen, S., Kamarck, T., & Mermelstein, R. (1983). A Global measure of perceived stress. *Journal of Health and Social Behavior*, 24(4), 385. <https://doi.org/10.2307/2136404>
- Creswell, J. D. (2017). Mindfulness interventions. *Annual Review of Psychology*, 68(1), 491–516. <https://doi.org/10.1146/annurev-psych-042716-051139>
- Davidson, R. J., & Kaszniak, A. W. (2015). Conceptual and methodological issues in research on mindfulness and meditation. *The American Psychologist*, 70(7), 581–592. <https://doi.org/10.1037/a0039512>
- Dawson, A. F., Brown, W. W., Anderson, J., Datta, B., Donald, J. N., Hong, K., Allan, S., Mole, T. B., Jones, P. B., & Galante, J. (2020). Mindfulness-based interventions for university students: A systematic review and meta-analysis of randomised controlled trials. *Applied Psychology: Health and Well-Being*, 12(2), 384–410. <https://doi.org/10.1111/aphw.12188>
- Dhillon, A., Sparkes, E., & Duarte, R. V. (2017). Mindfulness-based interventions during pregnancy: A systematic review and meta-analysis. *Mindfulness*, 8(6), 1421–1437. <https://doi.org/10.1007/s12671-017-0726-x>
- Dickenson, J., Berkman, E. T., Arch, J., & Lieberman, M. D. (2013). Neural correlates of focused attention during a brief mindfulness induction. *Social Cognitive and Affective Neuroscience*, 8(1), 40–47. <https://doi.org/10.1093/scan/nss030>
- Donald, J. N., Bradshaw, E. L., Ryan, R. M., Basarkod, G., Ciarrochi, J., Duineveld, J. J., Guo, J., & Sahdra, B. K. (2020). Mindfulness and its association with varied types of motivation: A systematic review and meta-analysis using self-determination theory. *Personality and Social Psychology Bulletin*, 46(7), 1121–1138. <https://doi.org/10.1177/0146167219896136>
- Donald, J. N., Sahdra, B. K., Van Zanden, B., Duineveld, J. J., Atkins, P. W. B., Marshall, S. L., & Ciarrochi, J. (2019). Does your mindfulness benefit others? A systematic review and meta-analysis of the link between mindfulness and pro-social behaviour. *British Journal of Psychology*, 110(1), 101–125. <https://doi.org/10.1111/bjop.12338>
- Eisenlohr-Moul, T. A., Peters, J. R., Pond, R. S., & DeWall, C. N. (2016). Both trait and state mindfulness predict lower aggressiveness via anger rumination: A multilevel mediation analysis. *Mindfulness*, 7(3), 713–726. <https://doi.org/10.1007/s12671-016-0508-x>
- Erisman, S. M., & Roemer, L. (2011). A preliminary investigation of the process of mindfulness. *Mindfulness*, 3(1), 30–43. <https://doi.org/10.1007/s12671-011-0078-x>
- Feldman, G., Greeson, J., & Senville, J. (2010). Differential effects of mindful breathing, progressive muscle relaxation, and loving-kindness meditation on decentering and negative reactions to repetitive thoughts. *Behaviour Research and Therapy*, 48(10), 1002–1011. <https://doi.org/10.1016/j.brat.2010.06.006>
- 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. <https://doi.org/10.1007/s10862-006-9035-8>
- Field, A., Miles, J., & Field, Z. (2012). *Discovering statistics using R*. SAGE Publications.
- Field, A. (2024). *Discovering statistics using IBM SPSS statistics*. Sage Publications Limited.
- Friese, M., & Hofmann, W. (2016). State mindfulness, self-regulation, and emotional experience in everyday life. *Motivation Science*, 2(1), 1–14. <https://doi.org/10.1037/mot0000027>
- Gavrilova, L., & Zawadzki, M. J. (2023). Mindfulness mechanisms in everyday life: Examining variance in acceptance, attention monitoring, decentering, self-compassion, and nonreactivity and their links to negative emotions among a workplace sample. *Cognition and Emotion*, 37(7), 1261–1271. <https://doi.org/10.1080/02699931.2023.2252960>
- Gethin, R. (2011). On some definitions of mindfulness. *Contemporary Buddhism*, 12(1), 263–279. <https://doi.org/10.1080/14639947.2011.564843>
- Gill, L.-N., Renault, R., Campbell, E., Rainville, P., & Khoury, B. (2020). Mindfulness induction and cognition: A systematic review and meta-analysis. *Consciousness and Cognition*, 84, 102991. <https://doi.org/10.1016/j.concog.2020.102991>

- Goldberg, S. B., Riordan, K. M., Sun, S., & Davidson, R. J. (2022). The empirical status of mindfulness-based interventions: A systematic review of 44 meta-analyses of randomized controlled trials. *Perspectives on Psychological Science*, 17(1), 108–130. <https://doi.org/10.1177/1745691620968771>
- Goldberg, S. B., Tucker, R. P., Greene, P. A., Davidson, R. J., Wampold, B. E., Kearney, D. J., & Simpson, T. L. (2018). Mindfulness-based interventions for psychiatric disorders: A systematic review and meta-analysis. *Clinical Psychology Review*, 59, 52–60. <https://doi.org/10.1016/j.cpr.2017.10.011>
- Goyal, M., Singh, S., Sibinga, E. M. S., Gould, N. F., Rowland-Seymour, A., Sharma, R., Berger, Z., Sleicher, D., Maron, D. D., Shihab, H. M., Ranasinghe, P. D., Linn, S., Saha, S., Bass, E. B., & Haythornthwaite, J. A. (2014). Meditation programs for psychological stress and well-being: A systematic review and meta-analysis. *JAMA Internal Medicine*, 174(3), 357–368. <https://doi.org/10.1001/jamainternmed.2013.13018>
- Gross, M., Raynes, S., Schooler, J. W., Guo, E., & Dobkins, K. (2024). When is a wandering mind unhappy? The role of thought valence. *Emotion*, 25(3), 671–682. <https://doi.org/10.1037/emo0001434>
- Grossman, P. (2008). On measuring mindfulness in psychosomatic and psychological research. *Journal of Psychosomatic Research*, 64(4), 405–408. <https://doi.org/10.1016/j.jpsychores.2008.02.001>
- Grossman, P. (2011). Defining mindfulness by how poorly I think I pay attention during everyday awareness and other intractable problems for psychology's (re) invention of mindfulness: Comment on Brown et al. (2011). *Psychological Assessment*, 23, 1034–1040; discussion 1041. <https://doi.org/10.1037/a0022713>
- 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. <https://doi.org/10.1080/14639947.2011.564841>
- Gu, J., Strauss, C., Crane, C., Barnhofer, T., Karl, A., Cavanagh, K., & Kuyken, W. (2016). Examining the factor structure of the 39-item and 15-item versions of the Five Facet Mindfulness Questionnaire before and after mindfulness-based cognitive therapy for people with recurrent depression. *Psychological Assessment*, 28(7), 791. <https://doi.org/10.1037/pas0000263>
- Guvendir, M. A., & Özkan, Y. Ö. (2022). Item removal strategies conducted in exploratory factor analysis: A comparative study. *International Journal of Assessment Tools in Education*, 9(1), 165–180. <https://doi.org/10.21449/ijate.827950>
- Hair, J. F., Black, W. C., Babin, B. J., & Anderson, R. E. (2019). *Multivariate data analysis* (8th ed.). Cengage Learning.
- Heppner, W. L., & Shirk, S. D. (2018). Mindful moments: A review of brief, low-intensity mindfulness meditation and induced mindful states. *Social and Personality Psychology Compass*, 12(12), e12424. <https://doi.org/10.1111/spc3.12424>
- Hu, L., & Bentler, P. M. (1999). Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. *Structural Equation Modeling: A Multidisciplinary Journal*, 6(1), 1–55. <https://doi.org/10.1080/10705519909540118>
- Johnson, S., Gur, R. M., David, Z., & Currier, E. (2013). One-Session mindfulness meditation: A randomized controlled study of effects on cognition and mood. *Mindfulness*, 6(1), 88–98. <https://doi.org/10.1007/s12671-013-0234-6>
- Kabat-Zinn, J. (1982). An outpatient program in behavioral medicine for chronic pain patients based on the practice of mindfulness meditation: Theoretical considerations and preliminary results. *General Hospital Psychiatry*, 4, 33–47. [https://doi.org/10.1016/0163-8343\(82\)90026-3](https://doi.org/10.1016/0163-8343(82)90026-3)
- Kabat-Zinn, J. (2003). Mindfulness-based interventions in context: past, present, and future. *Clinical Psychology: Science and Practice*, 10(2), 144–156. <https://doi.org/10.1093/clipsy.bpg016>
- Kahneman, D., Krueger, A. B., Schkade, D. A., Schwarz, N., & Stone, A. A. (2004). A survey method for characterizing daily life experience: The day reconstruction method. *Science*, 306(5702), 1776–1780. <https://doi.org/10.1126/science.1103572>
- Kakudji, B. K., Mwila, P. K., Burger, J. R., & Du Plessis, J. M. (2020). Epidemiological, clinical and diagnostic profile of breast cancer patients treated at Potchefstroom regional hospital, South Africa, 2012-2018: An open-cohort study. *The Pan African Medical Journal*, 36, 9. <https://doi.org/10.11604/pamj.2020.36.9.21180>
- Karl, J. A., & Fischer, R. (2020). Revisiting the five-facet structure of mindfulness. *Measurement Instruments for the Social Sciences*, 2, 1–16.
- Keng, S.-L., Smoski, M. J., & Robins, C. J. (2011). Effects of mindfulness on psychological health: A review of empirical studies. *Clinical Psychology Review*, 31(6), 1041–1056. <https://doi.org/10.1016/j.cpr.2011.04.006>
- Killingsworth, M. A., & Gilbert, D. T. (2010). A wandering mind is an unhappy mind. *Science*, 330(6006), 932–932.
- Kline, P. (2015). *A handbook of test construction (psychology revivals): Introduction to psychometric design*. Routledge.
- Lau, M. A., Bishop, S. R., Segal, Z. V., Buis, T., Anderson, N. D., Carlson, L., Shapiro, S., Carmody, J., Abbey, S., & Devins, G. (2006). The Toronto mindfulness scale: Development and validation. *Journal of Clinical Psychology*, 62(12), 1445–1467. <https://doi.org/10.1002/jclp.20326>
- Lecuona, O., García-Garzón, E., García-Rubio, C., & Rodríguez-Carvajal, R. (2020). A psychometric review and conceptual replication study of the Five Facets Mindfulness Questionnaire latent structure. *Assessment*, 27(5), 859–872. <https://doi.org/10.1177/1073191119873718>
- Li, C. H. (2016). Confirmatory factor analysis with ordinal data: Comparing robust maximum likelihood and diagonally weighted least squares. *Behavior Research Methods*, 48(3), 936–949. <https://doi.org/10.3758/s13428-015-0619-7>
- López, E., Jódar, R., Halty, L., & Rodríguez-Carvajal, R. (2022). Meditative State Scale (MSS): Psychometric validation and exploration of gamma-band correlates. *Mindfulness*, 13(10), 2558–2573. <https://doi.org/10.1007/s12671-022-01979-4>
- Ludwig, D. S., & Kabat-Zinn, J. (2008). Mindfulness in medicine. *JAMA*, 300(11), 1350–1352. <https://doi.org/10.1001/jama.300.11.1350>

- Lustig, J., Cardaciotto, L., Moon, S., & Spokas, M. (2024). Conceptualizations of mindfulness among experienced practitioners. *Mindfulness*, 15(4), 819–834. <https://doi.org/10.1007/s12671-024-02339-0>
- Mardia, K. V. (1970). Measures of multivariate skewness and kurtosis with applications. *Biometrika*, 57(3), 519–530. <https://doi.org/10.1093/biomet/57.3.519>
- Mattes, J. (2019). Systematic review and meta-analysis of correlates of FFMQ mindfulness facets. *Frontiers in Psychology*, 10, 2684. <https://www.frontiersin.org/articles/10.3389/fpsyg.2019.02684>
- Medvedev, O. N., Krägeloh, C. U., Siegert, R. J., & Singh, N. N. (2022). An introduction to assessment in mindfulness research. In O. N. Medvedev, C. U. Krägeloh, R. J. Siegert, & N. N. Singh (Eds.), *Handbook of assessment in mindfulness research* (pp. 1–11). Springer International Publishing. https://doi.org/10.1007/978-3-030-77644-2_1-1
- Medvedev, O. N., Norden, P. A., Krägeloh, C. U., & Siegert, R. J. (2018). Investigating unique contributions of dispositional mindfulness facets to depression, anxiety, and stress in general and student populations. *Mindfulness*, 9(6), 1757–1767. <https://doi.org/10.1007/s12671-018-0917-0>
- Mrazek, M., Smallwood, J., & Schooler, J. (2012). Mindfulness and mind-wandering: finding convergence through opposing constructs. *Emotion*, 12, 442–448. <https://doi.org/10.1037/a0026678>
- Navarrete, J., Fontana-McNally, M., Colomer-Carbonell, A., Sanabria-Mazo, J. P., Pinazo, D., Silvestre-López, A.-J., Lau, M., Tanay, G., García-Campayo, J., Demarzo, M., Soler, J., Cebolla, A., Feliu-Soler, A., & Luciano, J. V. (2023). The Toronto mindfulness scale and the state mindfulness scale: psychometric properties of the Spanish versions. *Frontiers in Psychology*, 14, 1212036. <https://doi.org/10.3389/fpsyg.2023.1212036>
- Neff, K. D., Tóth-Király, I., Knox, M. C., Kuchar, A., & Davidson, O. (2021). The development and validation of the State Self-Compassion Scale (Long- and Short Form). *Mindfulness*, 12(1), 121–140. <https://doi.org/10.1007/s12671-020-01505-4>
- Nilsson, H., & Kazemi, A. (2016). Reconciling and thematizing definitions of mindfulness: The big five of mindfulness. *Review of General Psychology*, 20(2), 183–193. <https://doi.org/10.1037/gpr0000074>
- Nunnally, J. C. (1967). *Psychometric theory* (pp. xiii, 640). McGraw-Hill.
- Pang, D., & Ruch, W. (2019). Scrutinizing the components of mindfulness: Insights from current, past, and non-meditators. *Mindfulness*, 10(3), 492–505. <https://doi.org/10.1007/s12671-018-0990-4>
- Park, T., Reilly-Spong, M., & Gross, C. R. (2013). Mindfulness: A systematic review of instruments to measure an emergent patient-reported outcome (PRO). *Quality of Life Research*, 22(10), 2639–2659. <https://doi.org/10.1007/s11136-013-0395-8>
- Petrocchi, N., & Ottaviani, C. (2016). Mindfulness facets distinctively predict depressive symptoms after two years: The mediating role of rumination. *Personality and Individual Differences*, 93, 92–96. <https://doi.org/10.1016/j.paid.2015.08.017>
- Pett, M. A., Pett, Lackey, N. R., & Sullivan, J. J. (2003). *Making sense of factor analysis: The use of factor analysis for instrument development in health care research*. SAGE.
- Poerio, G. L., Totterdell, P., & Miles, E. (2013). Mind-wandering and negative mood: Does one thing really lead to another? *Consciousness and Cognition*, 22(4), 1412–1421. <https://doi.org/10.1016/j.concog.2013.09.012>
- Preacher, K. J., & Hayes, A. F. (2008). Asymptotic and resampling strategies for assessing and comparing indirect effects in multiple mediator models. *Behavior Research Methods*, 40(3), 879–891.
- Preissner, C. E., Vilier, L., De Vries, H., & Oenema, A. (2024). Consistency between definitions and measurement of mindfulness in eating and physical activity behavior: A scoping review. *Health Psychology Review*, 18(3), 574–598. <https://doi.org/10.1080/17437199.2024.2310115>
- Prieto-Fidalgo, Á., Gómez-Odrizola, J., Royuela-Colomer, E., Orue, I., Fernández-González, L., Oñate, L., Cortazar, N., Iraurgi, I., & Calvete, E. (2021). Predictive associations of dispositional mindfulness facets with anxiety and depression: A meta-analytic structural equation modeling approach. *Mindfulness*, 13(1), 37–53. <https://doi.org/10.1007/s12671-021-01756-9>
- Quaglia, J. T., Brown, K., Lindsay, E., Creswell, J. D., & Goodman, R. J. (2015). From conceptualization to operationalization of mindfulness. In K. W. Brown, J. D. Creswell, & R. M. Ryan (Eds.), *Handbook of mindfulness: Theory, research and practice* (pp. 151–170). New York, NY: Guilford Press.
- R Core Team (2022). *R: A Language and environment for statistical computing*. R Foundation for Statistical Computing, Vienna, Austria. <https://www.R-project.org/>
- Raftery, A. E. (1995). Bayesian model selection in social research. *Sociological Methodology*, 25, 111–163. <https://doi.org/10.2307/271063>
- Raugh, I. M., Spilka, M., Luther, L., Suveg, C. M., & Strauss, G. P. (2023). Ecological momentary assessment of state fluctuations in mindfulness and symptoms in psychotic disorders. *Journal of Contextual Behavioral Science*, 29, 219–229. <https://doi.org/10.1016/j.jcbs.2023.08.002>
- Raynes, S., & Dobkins, K. (2025). Daily mindfulness components and state affect: A day reconstruction study. *Mindfulness*, 1–18. <https://doi.org/10.1007/s12671-025-02543-6>
- Robinson, M. D., & Clore, G. L. (2002a). Belief and feeling: Evidence for an accessibility model of emotional self-report. *Psychological Bulletin*, 128(6), 934–960. <https://doi.org/10.1037/0033-2909.128.6.934>
- Robinson, M. D., & Clore, G. L. (2002b). Episodic and semantic knowledge in emotional self-report: Evidence for two judgment processes. *Journal of Personality and Social Psychology*, 83(1), 198–215. <https://doi.org/10.1037/0022-3514.83.1.198>
- Rosseel, Y. (2012). lavaan: An R package for structural equation modeling. *Journal of Statistical Software*, 48(2), 1–36. <https://doi.org/10.18637/jss.v048.i02>
- Russ, S. L., Maruyama, G., Sease, T. B., & Jellema, S. (2017). Do early experiences matter? Development of an Early

- Meditation Hindrances Scale linked to novice meditators' intention to persist. *Psychology of Consciousness: Theory, Research, and Practice*, 4(3), 274–287. <https://doi.org/10.1037/cns0000129>
- Sauer, S., Walach, H., Schmidt, S., Hinterberger, T., Lynch Büssing, A., & Kohls, N. (2012). Assessment of mindfulness: review on state of the art. *Mindfulness*, 4, 3–17. <https://doi.org/10.1007/s12671-012-0122-5>
- Schlosser, M., Jones, R., Demnitz-King, H., & Marchant, N. L. (2022). Meditation experience is associated with lower levels of repetitive negative thinking: The key role of self-compassion. *Current Psychology*, 41(5), 3144–3155. <https://doi.org/10.1007/s12144-020-00839-5>
- Schumer, M. C., Lindsay, E. K., & Creswell, J. D. (2018). Brief Mindfulness training for negative affectivity: A systematic review and meta-analysis. *Journal of Consulting and Clinical Psychology*, 86(7), 569–583. <https://doi.org/10.1037/ccp0000324>
- Snippe, E., Bos, E., Ploeg, K., Sanderman, R., Fleer, J., & Schroevers, M. (2015). Time-series analysis of daily changes in mindfulness, repetitive thinking, and depressive symptoms during mindfulness-based treatment. *Mindfulness*, 6, 1053–1062. <https://doi.org/10.1007/s12671-014-0354-7>
- Solem, S., Thunes, S. S., Hjemdal, O., Hagen, R., & Wells, A. (2015). A metacognitive perspective on mindfulness: An empirical investigation. *BMC Psychology*, 3(1), 24. <https://doi.org/10.1186/s40359-015-0081-4>
- Spielberger, C. D. (1983). *State-trait anxiety inventory for adults* [dataset]. <https://doi.org/10.1037/t06496-000>
- Streiner, D. L. (2003). Starting at the beginning: An introduction to coefficient alpha and internal consistency. *Journal of Personality Assessment*, 80(1), 99–103. https://doi.org/10.1207/S15327752JPA8001_18
- Strickland, J. C., & Johnson, M. W. (2021). Rejecting impulsivity as a psychological construct: A theoretical, empirical, and sociocultural argument. *Psychological Review*, 128(2), 336–361. <https://doi.org/10.1037/rev0000263>
- Tan, L. B. G., Lo, B. C. Y., & Macrae, C. N. (2014). Brief mindfulness meditation improves mental state attribution and empathizing. *PLoS ONE*, 9(10), e110510. <https://doi.org/10.1371/journal.pone.0110510>
- Tanay, G., & Bernstein, A. (2013). State mindfulness scale (SMS): Development and initial validation. *Psychological Assessment*, 25. <https://doi.org/10.1037/a0034044>
- Thompson, B. L., & Waltz, J. (2007). Everyday mindfulness and mindfulness meditation: Overlapping constructs or not? *Personality and Individual Differences*, 43(7), 1875–1885. <https://doi.org/10.1016/j.paid.2007.06.017>
- Tran, U. S., Cebolla, A., Glück, T. M., Soler, J., Garcia-Campayo, J., & von Moy, T. (2014). The serenity of the meditating mind: a cross-cultural psychometric study on a two-factor higher order structure of mindfulness, its effects, and mechanisms related to mental health among experienced meditators. *PLoS ONE*, 9(10), e110192. <https://doi.org/10.1371/journal.pone.0110192>
- Tran, U. S., Glück, T. M., & Nader, I. W. (2013). Investigating the Five Facet Mindfulness Questionnaire (FFMQ): Construction of a short form and evidence of a two-factor higher order structure of mindfulness. *Journal of Clinical Psychology*, 69(9), 951–965. <https://doi.org/10.1002/jclp.21996>
- Truong, Q. C., Krägeloh, C. U., Siegert, R. J., Landon, J., & Medvedev, O. N. (2020). Applying generalizability theory to differentiate between trait and state in the Five Facet Mindfulness Questionnaire (FFMQ). *Mindfulness*, 11(4), 953–963. <https://doi.org/10.1007/s12671-020-01324-7>
- Ullrich-French, S., Cox, A., Cole, A., Rhoades Cooper, B., & Gotch, C. (2017). Initial validity evidence for the state mindfulness scale for physical activity with youth. *Measurement in Physical Education and Exercise Science*, 21(4), 177–189. <https://doi.org/10.1080/1091367X.2017.1321543>
- Ullrich-French, S., Cox, A. E., & Huong, C. (2021). The state mindfulness scale for physical activity 2: expanding the assessment of monitoring and acceptance. *Measurement in Physical Education and Exercise Science*, 26(2), 116–129. <https://doi.org/10.1080/1091367X.2021.1952207>
- Van Dam, N. T., van Vugt, M. K., Vago, D. R., Schmalzl, L., Saron, C. D., Olendzki, A., Meissner, T., Lazar, S. W., Kerr, C. E., Gorchov, J., Fox, K. C. R., Field, B. A., Britton, W. B., Brefczynski-Lewis, J. A., & Meyer, D. E. (2018). Mind the hype: A critical evaluation and prescriptive agenda for research on mindfulness and meditation. *Perspectives on Psychological Science*, 13(1), 36–61. <https://doi.org/10.1177/1745691617709589>
- Vinchurkar, S., Singh, D., Visweswaraiyah, N., & Nagendra, H. (2014). Immediate effects of cyclic meditation on state mindfulness in normal healthy volunteers: A controlled study. *Indian Journal of Positive Psychology*, 5, 400–403.
- Weijters, B., & Baumgartner, H. (2012). Misresponse to reversed and negated items in surveys: A review. *Journal of Marketing Research*, 49(5), 737–747. <https://doi.org/10.1509/jmr.11.0368>
- Weinstein, N., Brown, K. W., & Ryan, R. M. (2009). A multi-method examination of the effects of mindfulness on stress attribution, coping, and emotional well-being. *Journal of Research in Personality*, 43(3), 374–385. <https://doi.org/10.1016/j.jrp.2008.12.008>
- Williams, M., Honan, C., Skromanis, S., Sanderson, B., & Matthews, A. J. (2023). Psychological and attentional outcomes following acute mindfulness induction among high anxiety individuals: A systematic review and meta-analysis. *Journal of Psychiatric Research*, 170, 361–374. <https://doi.org/10.1016/j.jpsychires.2023.12.009>
- Zeidan, F., Emerson, N. M., Farris, S. R., Ray, J. N., Jung, Y., McHaffie, J. G., & Coghill, R. C. (2015). Mindfulness meditation-based pain relief employs different neural mechanisms than placebo and sham mindfulness meditation-induced analgesia. *Journal of Neuroscience*, 35(46), 15307–15325. <https://doi.org/10.1523/JNEUROSCI.2542-15.2015>