UNIVERSITY OF CALIFORNIA SAN DIEGO

Refining the Measurement and Analysis of State Mindfulness

A Dissertation submitted in partial satisfaction of the requirements
for the degree Doctor of Philosophy

in

Experimental Psychology

by

Stephen Raynes

Committee in charge:

Professor Karen Dobkins, Chair
Professor Janna Dickenson
Professor Michael McCullough
Professor Fadel Zeidan

2024

The Dissertation of Stephen Raynes is approved, and it is acceptable in quality and form for publication on microfilm and electronically.

University of California San Diego

2024

DEDICATION

To Arthur Thaddeus Raynes

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

LIST OF ABBREVIATIONS

State-FFMQ	State Four Facet Mindfulness Questionnaire
Trait-FFMQ	Five Facet Mindfulness Questionnaire
State-MAAS	State Mindful Attention and Awareness Scale
SMS	State Mindfulness Scale
TMS	Toronto Mindfulness Scale
MSMQ	Multidimensional State Mindfulness Questionnaire
MLM	Multilevel Model
ESM	Experience Sampling Method
DRM	Day Reconstruction Method

ACKNOWLEDGEMENTS

I often imagined that to write these acknowledgements, I would dedicate several days to creating a poetic expression of my heartfelt gratitude towards my support network, and articulating an interesting if not slightly pompous declaration about how my seemingly unique experiences over the past five years actually have ramifications that touch on the universality of some "deeper" aspects of the human experience. But priorities change. My time and attention have been grounded away from the clouds and towards several monumental upcoming events, including marrying my soulmate, having our first child, moving across the country, and starting a new job. Therefore, I would like to *briefly* express my appreciation to those who have helped me complete this dissertation: Professor Karen Dobkins for her expertise, mentorship, and patience as my advisor and the chair of my committee; my dissertation committee, department colleagues, and peers in the Human Experience and Awareness Lab for their invaluable advice; mentors from my previous educational programs for providing the inspiration for me to pursue higher education; my family and friends for their enduring support in all forms; and Hope Trout for providing a bottomless well of love, wisdom, and laughter.

Chapter 1 has been submitted for publication of the material as it may appear in Mindfulness, 2024, Raynes, Stephen; Dobkins, Karen, Springer. The dissertation author was the primary researcher and author of this paper.

Chapter 2 has been submitted for publication of the material as it may appear in Mindfulness, 2024, Raynes, Stephen; Dobkins, Karen, Springer. The dissertation author was the primary researcher and author of this paper.

Chapter 3 is currently under revision and is anticipated to appear in Emotion, 2024, Raynes, Stephen; Gross, Madeleine; Schooler, Jonathan; Guo, Evie; Dobkins, Karen, APA PsycArticles. The dissertation author was the primary researcher and author of this papers.

VITA

2015 Bachelor of Arts in Psychology (major) and Business Management (special concentration), Columbia University

2018 Master of Research in Developmental Neuroscience and Psychopathology, University College London and Yale University (joint degree program)

2021 Master of Arts in Experimental Psychology, University of California San Diego

2024 Doctor of Philosophy in Experimental Psychology, University of California San Diego

FIELD OF STUDY

Major Field: Experimental Psychology
Studies in Social Psychology
Professor Karen Dobkins

ABSTRACT OF THE DISSERTATION

Refining the Measurement and Analysis of State Mindfulness

by

Stephen Raynes

Doctor of Philosophy in Experimental Psychology

University of California San Diego, 2024

Professor Karen Dobkins, Chair

Research examining the benefits of mindfulness have expanded exponentially over the past few decades, yet the field faces criticism for lacking methodological rigor, particularly in terms of how mindfulness is measured. Most of the effort towards developing psychometrically sound self-report measures has been made regarding trait mindfulness. However, there is good reason to develop valid and reliable state measures since mindfulness is often referred to as a momentary state that is experienced during formal practices like meditation, as well as

throughout instances of daily life, to varying degrees. Although a handful of existing state mindfulness measures have merit, I argue the need for a new measure that is brief, multidimensional, accessible to the general population, and versatile to various experimental designs. With these motivations in mind, the goal of this dissertation is to improve the methodological rigor by which researchers can analyze the benefits of mindfulness. In Chapter 1, I create and validate a new measure of state mindfulness, the State Four Facet Mindfulness Questionnaire ("state-FFMQ"), fashioned after the most widely used multidimensional measure of trait mindfulness. In an experimental setting, the state-FFMQ provides evidence of convergent, construct, predictive, and incremental validity when referencing a formal mindfulness practice. In Chapter 2, I employ the Day Reconstruction Method to show that in a naturalistic setting the state-FFMQ demonstrates these same forms of validity when referencing daily experiences of mindfulness. In both chapters, I also investigate the empirical question of which facet(s) of state mindfulness best predicts state affect and show that not all facets of mindfulness are alike. In Chapter 3, I replicate a widely-cited study that looks at the association of state mindfulness and state affect and expand on their results by quantifying the impact of a rarely evaluated yet significant confounding variable. In sum, this dissertation provides evidence that a newly created state mindfulness measure is psychometrically sound; clarifies the relative importance of each facet of state mindfulness in predicting state affect within experimental and naturalistic settings; and determines how much of the relationship between state mindfulness and state affect can be accounted for by a confounder.

INTRODUCTION

Research examining the benefits of mindfulness have expanded exponentially over the past few decades, yet the field faces criticism for lacking methodological rigor, particularly in terms of how mindfulness is measured. Since mindfulness is an inherently internal experience, it is difficult to detect with objective measures; therefore, most studies rely on self-report measures. Unfortunately, the development of valid and reliable self-report measures of mindfulness has proven to be a challenging task, with little consensus in the literature as to which measures ought to be used. Most of the effort towards developing psychometrically sound self-report measures has been made with regard to trait mindfulness. However, there is good reason to develop valid and reliable state measures since mindfulness is often referred to as a momentary state that is experienced during formal practices like meditation, as well as throughout instances of daily life, to varying degrees. Although a handful of existing state mindfulness measures have merit, there is a need for a new state measure that is brief, multidimensional, accessible to the general population, and versatile to various experimental designs. With these motivations in mind, the goal of this dissertation is to improve the methodological rigor by which researchers can analyze the benefits of mindfulness.

Chapter 1 includes two pre-registered studies that together create and validate a new measure of state mindfulness, the State Four Facet Mindfulness Questionnaire ("state-FFMQ"), fashioned after the most widely used multidimensional measure of trait mindfulness. In these studies, the state-FFMQ is measured in reference to an immediately preceding 20-minute mindfulness meditation session. Furthermore, the extent to which each facet of state mindfulness uniquely predicts the benefit of the meditation session via reductions in state anxiety and stress is investigated.

Chapter 2 includes one pre-registered study wherein the state-FFMQ is administered via an intensive longitudinal design (i.e. the Day Reconstruction Method) to capture the within-person variation in state mindfulness during recalled experiences from the previous day. Similar to Chapter 1, this chapter aims to examine the validity of the state-FFMQ, and to measure the extent to which each facet of state mindfulness uniquely predicts momentary happiness during daily life. We also explore whether there is any long-term psychological benefit to completing the Day Reconstruction Method per se.

In Chapters 1 and 2, one unexpected finding is that the rarely measured construct of "valence" plays an important role as a covariate in predictive models. In Chapter 1, the self-reported valence of one's thoughts during a meditation uniquely predicts the benefits of a meditation session and does not share variance with state mindfulness. In Chapter 2, the self-reported valence of the activity one is engaged in is uniquely predictive of momentary happiness and does share variance with state mindfulness. Given these interesting results, Chapter 3 includes one pre-registered study that first replicates a widely-cited study examining the association between state mindfulness (measured unidimensionally via attention state) and momentary mood in daily life, and then expands on their results by quantifying the proportion of that association accounted for by valence (i.e. the self-reported valence of one's thoughts during daily life experiences).

Chapter 1 Initial Validation of the State Four Facet Mindfulness Questionnaire

Abstract

The current research aimed to provide initial psychometric validation of a new multifaceted mindfulness questionnaire (referred to as the "state-FFMQ") - 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 studies, undergraduate students listened to a 20-minute mindfulness meditation audio recording, then answered questions, including the state-FFMQ, pertaining to their experience during the recording. In Study 2, participants additionally listened to a 20-minute control audio recording. The state-FFMQ 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-FFMQ 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-FFMQ, and its short-form, showed good internal consistency as well as convergent and predictive validity. In addition, it was found that some facets, more than others, predicted momentary well-being. The validity of the state-FFMQ shows that it can be used to measure multiple facets of state mindfulness across a variety of situations.

General 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 1970's, 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 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 Buddhists 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- vs. 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: *Awareness* (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.

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 set up 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 timepoints (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 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 number of citations), by addressing the 1) the framework used to guide the creation of the scale, 2) the "referent experience" (e.g., a single experimental manipulation vs. an ILD) the scale was tested with, 3) multi-dimensionality of the scale, and 4) how the scale was validated (summarized in Table 1.1). 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 fourteen consecutive days), though the scale has since been used in other contexts such as in reference to an 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 1) 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 2) 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), and the scale has since been used in other contexts such as in reference to experimental tasks (e.g., watching a film clip; Erisman & Roemer, 2011), though it would not be applicable in other contexts such as daily activities for the general population. 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 1) 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 2) 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 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: 1) "Present-moment attention" (e.g., "I focused my attention on the present moment"), 2) "Acting with awareness" (e.g., "I sometimes did not stay focused on what was happening in the present", reverse scored), and 3) "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 - 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; Friese & Hofmann, 2016; Gavrilova & Zawadzki, 2023; Raugh et al., 2023; Snippe et al., 2015). Critically, however, none of these studies psychometrically validated their state adaptations, which is the goal of the current study.

Table 1.1: Validated State Mindfulness Measures

Measure Name	Dimensions	Theoretical Basis	Development Article	Citations[a]
Mindful Awareness and Attention Scale (state-	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.
[a]Google Scholar citations as of February 7, 2024. [b]This includes citations for the trait-MAAS, which was developed in the same article and thus cannot be separated.

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-FFMQ". Broadly speaking, an ideal state mindfulness measure would 1) be brief (noting that the 21-item SMS is too long to be used practically in an ILD design), 2) use accessible language (noting that the TMS may not be relevant to meditation-naive participants, and the MSMQ has not been validated in English), 3) 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), 4) include both positive and reverse coded items within a dimension to boost the measure's validity (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 5) 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. The final EFA structure was 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 KIMS (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 wellbeing. 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-FFMQ: construct validity (testing whether a measure designed to assess a particular construct is actually measuring that construct in an expected way), 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), incremental validity (testing whether the predictive value of the measure remains robust after accounting for covariates that are related to the measure).

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-FFMQ"). We hypothesized that multiple facets would emerge from the EFA results; but given the critiques of 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

Participants. Participants were undergraduate students recruited in 2022 through the UCSD SONA 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 5-10 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. *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 Appendix C). 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 limits 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 - 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%), and Mixed (7.30%). The remainder of participants were in one of several low frequency categories that together added up to 7.08%.

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 Appendix A 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-FFMQ. To create the state-FFMQ, 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, thereby controlling for response biases and increasing the validity of the measure (a commonly cited issue with the trait-FFMQ, e.g., Grossman & Van Dam, 2011). 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 63[1] items to be tested in our study (see Appendix B).

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 1-5 with a resolution of 0.1 and with

[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.

three labels: Not at all (1), Moderately (3), Completely/Entirely (5), 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 minutes to 60+ minutes. 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-.7 are mediocre, .7-.8 good, .8-.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.

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 fit 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 ($\chi 2$), 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 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 - 0.44), excess kurtosis (range = -1.25 - 0.30), and bivariate correlation values with other items (range = 0.10 - 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, X^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 1) resulted in an acceptable determinant value, 2) retained comprehensive coverage of all content areas (including heterogeneous content areas within a factor) covered in the trait-FFMQ, and 3) 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: 1) 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), 2) had been selected for use in a previously validated short-form of the trait-FFMQ (Gu et al., 2016), and 3) required minimal content modification when adapted from the trait-FFMQ, and 4) 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 Appendix B 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, X^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) 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 1.2 (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, Awareness 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 1.2 (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 - .89 (MacDonald's ω). Furthermore, the internal consistency of all items together was high ($\omega = .87$).

Table 1.2: Exploratory Factor Solution of the state-FFMQ

Item	Describing	Awareness	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
	Factor Correlation Coefficients			
Describing	-	.24	.15	.31
Awareness		-	.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. 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.

Table 1.2 (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-FFMQ 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 factory 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-FFMQ)* as well as a shortened version with 12-items (which we refer to as the *short-form state-FFMQ*, explained below). Second, data from Study 2 allowed us to further test the validity of the full- and short-form state-FFMQ, by testing for convergent, predictive, incremental, and construct validity.

Method

Participants. The participant pool and eligibility criteria were the same as described in Study 1.

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-FFMQ (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 (see Procedure below) 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, see Protocol, below) to not answering the survey questions honestly, or not fully engaging in the relaxation exercise, respectively (see Appendix C). *Fifth*, 4 participants were excluded for failing to have a total state-FFMQ score within +/- 3 standard deviations of the mean total 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 - 46 years (M = 20.45, 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%), and Mixed (9.50%). The remainder of participants were in one of several low frequency categories that together added up to 4.13%.

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 = .08; 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 (see Procedure below) 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%), and Mixed (9.32%). The remainder of participants were in one of several low frequency categories that together added up to 4.18%.

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 Appendix A 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). After the audio recording they completed the following measures in this order: the state-FFMQ, the two other state mindfulness surveys in randomized order, the two state affect measures in randomized order, the clarify 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-FFMQ. 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 $\alpha = .84 - .86$ 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 $\alpha = .90 - .93$ for the total score, and $\alpha = .79 - .90$ 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, 2012). 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 $\alpha = .93 - .94$ 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., n.d.; Weinstein et al., 2009). These three in-house items showed acceptable internal consistency with Cronbach’s alpha coefficients ranging from α = .82 - .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). There are two reasons to believe this Thought Valence metric is reliable. First, on an additional validity check in which participants were 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)), the mean ratings were relatively high (Meditation $M = 4.14$, $SD = 0.96$; Control $M = 4.02$, $SD = 0.94$). Second, we found in our tests for predictive validity that Thought Valence does, in fact, predict state affect (see Results), which indicates that the Thought Valence metric is sufficiently reliable. The wording of this in-house item was inspired by a previous study in our lab (Gross et al., under review) that found significant associations between Thought Valence in daily life, mindful attention, and mood.

Clarity of Instructions. Since guided meditations may seem esoteric or inaccessible to novice meditators, we include an additional question asked in Feldman (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 are available in the literature, and since it is the basis for the state-FFMQ. The scale showed acceptable internal consistency with Cronbach's alpha coefficients ranging from α = .90 - .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 α = .86 - .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 heterogenous 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 Appendix C 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 - .30 are weak effects, .30 - .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 ($\eta 2$) 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-FFMQ from Study 1 was relatively lengthy compared to most state mindfulness measures. We ultimately wanted the state-FFMQ 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-FFMQ using the same sample as the full-form (as has been done in the development of other measures, e.g., 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-FFMQ. To test this, all of the following analyses were conducted on both the full (18-item) and short (12-item) state-FFMQ.

The main goal of the CFA was to confirm the accuracy of the four-factor structure of the state-FFMQ that emerged from Study 1 through EFA. Prior to conducting the CFA, we first calculated bivariate correlations amongst the state-FFMQ 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-FFMQ 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-FFMQ 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 1) 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 2) 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 ($\chi 2$; $p > .05$), 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-FFMQ 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-FFMQ 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, 6–10 as strong evidence; 2–6 as positive evidence; and 0–2 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-FFMQ. 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-FFMQ were comparable to the short-form and are available in Appendix E). 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-FFMQ 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-FFMQ facets, although the Awareness component of the state-FFMQ 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-FFMQ'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 (e.g., state Awareness and trait Awareness) should be more positively correlated than non-aligned facets (e.g., state Awareness 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 show likewise show only moderate correlations.

Predictive Validity. Here, we tested whether any (or all) of the facets of the state-FFMQ 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-FFMQ facets were entered simultaneously as predictor variables. Because our state affect measures are "negative" states, we expected a negative relationship between the state-FFMQ 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-invention score for the state-FFMQ. 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-FFMQ items, as they are in the process of development. Second, had we attempted to obtain pre-intervention scores for the state-FFMQ items, it would have been difficult to standardize the pre-intervention reference point (e.g., responding in reference to what the participant was doing just before the study began). 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 Awareness and Nonjudging would most strongly predict state affect, followed by Describing, and that Observing will 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).

Incremental Validity. Given that we found evidence for predictive validity (see Results), in the next step, we tested whether that the strength of the relationship between state-FFMQ and state affect would be altered when other two variables were included in the model, namely, trait-FFMQ and Thought Valence. The inclusion of the trait-FFMQ weighed in on construct validity (see below), while the inclusion of Thought Valence weighed in on potential underlying mechanisms regarding the beneficial effects of mindfulness (see *Discussion).*

There are two reasons to include covariates in a model. First, if a covariate is strongly correlated with the dependent variable and weakly (or not) correlated with the state-FFMQ, its inclusion in the model can reduce variance in the dependent variable that would otherwise be unaccounted for, thereby enhancing the significant effects of the state-FFMQ. Second, if a covariate is strongly correlated with *both* the dependent variable as well as the state-FFMQ, its inclusion in the model can potentially account for the relationship between the state-FFMQ and the dependent variable. In this case, we would want to ensure that the relationship observed between the state-FFMQ and the dependent variable (in the original predictive validity analysis, *above*) was not *entirely* accounted for by this covariate. We confirmed whether the trait-FFMQ total score and Thought Valence had the potential to behave in either the first or second way by looking at bivariate relationships between the state-FFMQ facet scores, and the two covariates, and between the two covariates and the dependent variables (separately). We further confirmed whether a potential covariate interacted with the state-FFMQ 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-FFMQ was behaving in a way that was consistent with its intended design of 1) being a “State” vs. Trait” measure, as well as 2) detecting “mindful” states. Evidence that the state-FFMQ was behaving in a state-like way was addressed in two ways. First, when testing for incremental validity of the state-FFMQ (above), we reasoned that if the relationship between the state-FFMQ and state affect remained strong while accounting for the effects of the trait-FFMQ, this would demonstrate that the new state-FFMQ is not masquerading as a trait measure. Second, data from the two times points over which the state-FFMQ 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-FFMQ 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): PSS, trait-FFMQ, state-FFMQ, 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-FFMQ detects *mindful* states, we conducted a test of “discriminant sensitivity” wherein state-FFMQ scores were expected to be higher after the Meditation vs. the Control condition. This was implemented using a mixed ANOVA with the state-FFMQ 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.

Results

CFA Analysis. Descriptive statistics and bivariate correlations for the state-FFMQ facets are presented in Table 1.3. 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-FFMQ.

In the full-form state-FFMQ, 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-FFMQ demonstrated a similar pattern of results, with one additional insignificant relationship between Awareness and Nonjudging (*r*(240)= .09).

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

Table 1.3: Descriptive Statistics, Internal Consistency, and Bivariate Correlations of the state-FFMQ

Facet	*M (SD)*	α	ω	1	2	3	4
1. Awareness	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-FFMQ are presented left of the dash, and for the short-form state-FFMQ on the right of the dash. All correlations had 240 degrees of freedom.
* indicates $p < .05$. ** indicates $p < .01$.

Preliminary analyses showed that all 18 items of the state-FFMQ were approximately normally distributed, as assessed by levels of skewness (range -0.71 - 0.23) and excess kurtosis (range -1.12 - 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-FFMQ. 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-FFMQ, the four-factor nonhierarchical, hierarchical, and bifactor models demonstrated acceptable fit indices (RMSEA < .08; CFI/TLI > .90, see Table 1.4) 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-FFMQ, the four-factor nonhierarchical model fit better than the four-factor hierarchical model (full ΔBIC = 1.48; short ΔBIC = 15.95), and the four-factor hierarchical model fit better than the four-factor bifactor model (full ΔBIC = 42.86; short ΔBIC = 36.75). Though differences were more obvious in the short than the full state-FFMQ, 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, $p < .001$, RMSEA = 0.06 (90% CI 0.05 - 0.08), SRMR = 0.07, TLI = 0.91, CFI = 0.93) and short-form ($\chi 2$ (48) = 88.14, $p <.001$, RMSEA = 0.06 (90% CI 0.04 - 0.08), SRMR = 0.05, TLI = 0.94, CFI = 0.95) state-FFMQ.

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

Table 1.4: Model Fit and Comparison Indices

	Model Fit Indices					Model Comparison Indices	
Model	**RMSEA [90% CI]**	**SRMR**	**TLI**	**CFI**	**$\chi 2$ (*df*)**	**BIC**	**AIC**
Full-form state-FFMQ							
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
Short-form state-FFMQ							
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; $\chi 2$ = Scaled Chi-Square Statistic; df = degrees of freedom; BIC = Bayesian Information Criterion; AIC = Akaike information criterion. All $\chi 2$ statistics were significant at $p < .001$.

As the four-factor nonhierarchical model of the state-FFMQ was the best-fitting model, Table 1.5 displays summary statistics for this model only[2]. For the short-form state-FFMQ, 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-FFMQ, 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-FFMQ.

[2] Summary statistics of all models are available by request from the authors

Table 1.5: Confirmatory Factor Analysis of the Full and Short Four-Factor Nonhierarchical Models

Factor	Item	Loading
Awareness		
	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. Items retained in the short-form are bolded. Loadings for the full-form state-FFMQ are in parentheses.

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 - 0.77) and short (0.42 - 0.93) state-FFMQ.

Further Validation Analyses.

Convergent Validity. Convergent validity was tested by investigating whether the state-FFMQ 1) shares variance with other state mindfulness measures and 2) shares variance with the trait-FFMQ, moreso for aligned vs. misaligned facets. Bivariate correlations of all state and trait mindfulness[3] measures are presented in Table 1.6. As predicted, there was a strong correlation between the state-MAAS and the Awareness facet of the state-FFMQ ($r(309) = .69$), as well as some smaller correlations between the state-MAAS and the other three facets. The SMS subscales were somewhat strongly 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-FFMQ were weakly to moderately correlated with one another (*r's*(309) = .15 - .39), with the exception of one insignificant relationship between Observing and Nonjudging ($r(309) = .10$).

[3] Note that while we used the trait scores from Part 2, the results were similar if using the trait scores from Part 4.

Table 1.6: Convergent Validity of the Short-form state-FFMQ

Variable	1	2	3	4	5	6	7	8	9	10	11	12
1. State Awareness												
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 Awareness	**.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-FFMQ 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-FFMQ Total.
$^{*}p < .05$. $^{**}p < .01$.

Predictive Validity. Predictive validity was tested by asking if any (or all) facets of the state-FFMQ predicted another state two 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 in 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-FFMQ. The results are summarized in Table 1.7 - Table 1.8 (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, $\eta2$ = .23 - .35), which itself provides evidence for the reliability of the measures. With respect to state-FFMQ facets, all but one (negatively) predicted State Anxiety/State Stress. Specifically, Awareness and Nonjudging were significant predictors with medium effect sizes (*p*'s < .001, $\eta2$ = .04 - .05), followed closely by Observing with slightly smaller effect sizes (*p*'s = .002 - <.001, $\eta2$ = .03 - .04), while Describing was not significant in either model (*p*'s = .19 - .27).

Incremental Validity. 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-FFMQ facets) and the dependent variables (i.e., state affect). Specifically, the trait-FFMQ Total score positively correlated with all state-FFMQ 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-FFMQ 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 incremental analyses, we also ensured that neither covariate interacted with any state-FFMQ facet in the models.

As shown in Tables 7A - 7B (right panels), there were no notable differences in the beta coefficient, *p*-value, or effect size in any state-FFMQ facet after accounting for trait-FFMQ and Thought Valence. In addition, for both dependent measures, the trait-FFMQ Total failed to reach significance (*p's* = .20 - .92), while Thought Valence showed significant effect sizes that were on par with those of the state-FFMQ facets (*p's* < .001, η2 = .04 - .05). The robustness of the state-FFMQ's predictive validity when trait-FFMQ was included in the model is discussed below (see *Construct Validity*). The finding that the inclusion of Thought Valence did not lessen the relationship between state-FFMQ and state affect indicates that the relationship is not accounted for by Thought Valence (see *Discussion).*

Table 1.7: Multiple Linear Regression Results Predicting State Anxiety After a Single Meditation Session

Variable	Predictive Model				Incremental Model			
	B	*SE*	*p*	η2	*B*	*SE*	*p*	η2
(Intercept)	40.65	2.62	<.001		40.22	3.08	<.001	
Pre-Score	0.36	0.03	**<.001**	.35 [.27, .42]	0.37	0.03	**<.001**	.35 [.27, .42]
Awareness	-1.22	0.33	**<.001**	.04 [.01, .10]	-0.99	0.33	**.003**	.03 [.00, .07]
Describing	-0.47	0.36	.19	.006 [.00, .03]	-0.54	0.36	.13	.007 [.00, .04]
Nonjudging	-1.49	0.37	**<.001**	.05 [.01, .11]	-1.47	0.37	**<.001**	.05 [.01, .10]
Observing	-1.44	0.45	**.002**	.03 [.00, .08]	-1.37	0.45	**.002**	.03 [.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. η2 represents partial eta-squared. Numbers in brackets represent 95% confidence intervals. Collinearity as assessed with VIF were well within acceptable limits for each model. Predictors were entered into analysis simultaneously. Significant effects are bolded.

Table 1.8: Multiple Linear Regression Results Predicting State Stress After a Single Meditation Session

Variable	Predictive Model				Incremental Model			
	B	*SE*	*p*	η2	*B*	*SE*	*p*	η2
(Intercept)	9.54	0.86	<.001		10.11	1.00	<.001	
Pre-Score	0.37	0.04	**<.001**	.23 [.15, .31]	0.37	0.04	**<.001**	.23 [.16, .31]
Awareness	-0.54	0.13	**<.001**	.05 [.01, .11]	-0.42	0.13	**.002**	.03 [.00, .08]
Describing	-0.16	0.14	.27	.004 [.00, .03]	-0.14	0.14	.34	.003 [.00, .03]
Nonjudging	-0.51	0.15	**<.001**	.04 [.01, .09]	-0.45	0.15	**.002**	.03 [.00, .08]
Observing	-0.67	0.18	**<.001**	.04 [.01, .10]	-0.61	0.18	**.001**	.04 [.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. η2 represents partial eta-squared. Numbers in brackets represent 95% confidence intervals. Collinearity as assessed with VIF were well within acceptable limits for each model. Predictors were entered into analysis simultaneously. Significant effects are bolded.

Construct Validity.

State versus Trait aspects. The incremental validity analysis, above, provided one source of evidence that the state-FFMQ behaves in a state- vs. trait-like fashion. This is because the relationship between state-FFMQ 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-FFMQ was to compare test-retest reliability between the state-FFMQ and the trait-FFMQ. As shown in Table 1.9, the stability of the state-FFMQ 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.

Table 1.9: Test-Retest Reliability of State and Trait Measures

Measure	Correlation Between Timepoints
Chronic Stress	.83
Trait-FFMQ Total	.85
Trait Awareness	.82
Trait Describing	.83
Trait Nonjudging	.81
Trait Observing	.77
Trait Nonreactivity	.77
State-FFMQ Total	.62
State Awareness	.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

Discriminant sensitivity to mindful states. Here, we tested construct validity of the state-FFMQ by asking if it detects mindful states. If so, state-FFMQ 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 1.10 for current and past meditators).

Table 1.10: Results of Meditation Experience Items for Current and Past Meditators

	Current	**Past**	**Overall**
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	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)

[a] Participants 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).

The results from a mixed-ANOVA, in which the dependent variable was total state-FFMQ 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, η^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-FFMQ facet score as the dependent variable.

In sum, we did not find evidence for discriminant sensitivity of the state-FFMQ, 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 Appendix D. As we return to in 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-FFMQ 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-FFMQ 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-FFMQ was found to demonstrate convergent, predictive, and incremental validity. Results of construct validity were varied; whereas we found substantial evidence that the state-FFMQ acts more like a state than a trait, we could not yet confirm its discriminant sensitivity to detecting mindful states.

Chapter 1 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-FFMQ), 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-FFMQ is short (12 items for the short-form), multi-dimensional (4 facets), contains reversed- (as well as positive-) scored items within each facet, 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; determining which state mindfulness facets best predict state affect and asking whether

Thought Valence accounts for the relationship between state mindfulness and 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-FFMQ'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 its 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 factor, 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 have been conceptual criticisms 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.

Null Results for Discriminant Sensitivity. The other unexpected result was a null finding in an analysis of the state-FFMQ'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 Appendix D 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-FFMQ 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).

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 affect (see Ullrich-French et al., 2021). Note that in our sister paper (Raynes & Dobkins, under review), we ask this question in an ILD design. In the current study, we found unique contributions of Awareness 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). More broadly, the relative contribution of each facet likely varies based not only on the experimental context (e.g., mediation type), but also 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).

Does Thought Valence account for the Relationship between State Mindfulness and State Affect? Another empirical question that can be addressed from the current data is whether Thought Valence accounts for (and potentially mediates) the relationship between state mindfulness and state affect, which is rarely considered in meditation research. As might be expected, the current study found a moderate (and negative) bivariate correlation between Thought Valence and state affect. Because we also found significantly positive bivariate correlations between Thought Valence and the state-FFMQ facets, i.e., thoughts were more positive for participants who experienced higher levels of state mindfulness during the meditation, this leaves open the possibility that the negative relationship between state mindfulness and state affect is driven not by mindfulness itself, but instead, by mindfulness states resulting in more positive thoughts, which in turn improve state affect[4]. The results of the current study suggest this is not the case, since the relationship between state mindfulness and state affect was unaffected by the inclusion of Thought Valence in the model.

[4] Of course, one can also imagine that state mindfulness accounts for the observed relationship between Thought Valence and state affect, however, this was not tested (or pre-registered) in the current study.

In sum, the current findings suggest that state mindfulness and thought valence have unique (and independent) influences on state affect. Although the impact of thought valence on the association between *mind wandering* and state affect has been fairly well-studied (e.g., Banks et al., 2016; Mills et al., 2021; Poerio et al., 2013; Welz et al., 2018) there is a notable scarcity of similar investigations focusing on mindfulness, rather than mind wandering. Therefore, it is difficult to contextualize our findings in the broader mindfulness literature. However, another study (Gross et al., under review) employing repeated measurement of daily life experiences found that Thought Valence was more positive in moments when participants were in a more mindful state (in line with the current research); however, that study also found that Thought Valence *did* largely account for the relationship between state mindfulness and state affect. While the inconsistent findings between the current and previous studies are likely attributable to differing measurement scales, study designs, and analytical approaches, further research will be needed to elucidate the effects of Thought Valence on state mindfulness in various contexts.

Limitations. On a final note, we acknowledge some general limitations. As with any scale development article, creating the state-FFMQ 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-FFMQ 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 Appendix B and explore alternative approaches to developing a state-FFMQ. 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-FFMQ was developed using traditional psychometric methods (e.g., factor analysis, validity testing, etc.). 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-FFMQ 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 that 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-FFMQ 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 of 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-FFMQ 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.

Acknowledgements

Chapter 1 has been submitted for publication of the material as it may appear in Mindfulness, 2024, Raynes, Stephen; Dobkins, Karen, Springer. The dissertation author was the primary researcher and author of this paper.

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

Banks, J. B., Welhaf, M. S., Hood, A. V. B., Boals, A., & Tartar, J. L. (2016). Examining the role of emotional valence of mind wandering: All mind wandering is not equal. *Consciousness and Cognition*, *43*, 167–176. https://doi.org/10.1016/j.concog.2016.06.003

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., Speca, 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*. https://doi.org/10.1007/s12671-016-0651-4

Bondi, T. E. (2021). Mechanisms of Meditation: Investigating the Components and Covariates of a Single Session of Meditation [UC San Diego]. https://escholarship.org/uc/item/3v927741

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*. 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., & Baitmangalkar, 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*. 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 prosocial 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.

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*, 1–11. 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

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

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 Chornic 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

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

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

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

López, E., Jódar, R., Halty, L., & Rodríguez-Carvajal, R. (2022). Meditative State Scale (MSS): Psychometric Validation and Exploration of Gamma-Band Correlates. *Mindfulness*. 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*. 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*. 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

Mills, C., Porter, A. R., Andrews-Hanna, J. R., Christoff, K., & Colby, A. (2021). How task-unrelated and freely moving thought relate to affect: Evidence for dissociable patterns in everyday life. *Emotion*, *21*(5), 1029–1040. https://doi.org/10.1037/emo0000849

Mrazek, M., Smallwood, J., & Schooler, J. (2012). Mindfulness and Mind-Wandering: Finding Convergence Through Opposing Constructs. *Emotion (Washington, D.C.)*, *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

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*, 1–25. https://doi.org/10.1080/17437199.2024.2310115

Prieto-Fidalgo, Á., Gómez-Odriozola, 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*. 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 *Handbook of mindfulness: Theory, research, and practice* (pp. 151–170).

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

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*, *048*(i02). https://econpapers.repec.org/article/jssjstsof/v_3a048_3ai02.htm

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*. 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*. 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

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*, 1–14. 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., Visweswaraiah, 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.

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

Welz, A., Reinhard, I., Alpers, G. W., & Kuehner, C. (2018). Happy Thoughts: Mind Wandering Affects Mood in Daily Life. *Mindfulness*, *9*(1), 332–343. https://doi.org/10.1007/s12671-017-0778-y

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*. 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

Appendices

Appendix A

Condition Transcripts

Meditation condition

(5 second pause)
Welcome, I will be leading you through this exercise today. To begin, please close your eyes and relax, and let out a long, slow exhale.
(7s)
Relax all the muscles in your body, and take a few long, slow breaths through your nose.
(15s)
Bring your attention to the sensation of your breath. Slowly inhale and exhale through the nose, allowing each breath to flow more deeply.
(20s)
Find an easy, slow rhythm in your breath. Notice the air moving in and out of your nostrils, the air going up and down the back of your throat, expanding and contracting your chest and abdomen.
(30s)
[2:00] When you get distracted, either by your thoughts or feelings, simply notice the thought or feeling, and return your focus back to the breath.
(30s)
As thoughts come and go, try not to judge them as good or bad. Simply notice the thoughts move in and out of your mind, let them pass as you draw your awareness back to your breath.
(45s)
[4:00] Witness the flow of your breath, the pauses and spaciousness of the in breath and the out breath. Notice the filling of your lungs and belly as you breathe in, the emptying as you breathe out.
(1 min)
When your mind wanders, gently bring it back to the breath.
(45s)
Just focus on the sensations of your breath.
(1 min)
[7:45] Continue to notice passing thoughts with curiosity and compassion, let it go without judgement, and bring your attention back to the natural flow of your breath.
(1 min)
Let thoughts pass like clouds passing over a clear sky, and focus on the gentle flow of your breath.
(1 min)
[10:30] If your mind wanders, gently bring it back to the sensations of your breath.
(1 min)
Notice the steady expansion as you breathe in, the easy release as you breathe out.
(1 min)
[13:00] Continue to practice this on your own for the next few minutes. When your mind wanders, acknowledge the thought with compassion, and gently bring your focus back to the breath.
(2 min, 30s)
Just notice your breath. Allowing thoughts to pass without judgement. Continue to practice this in silence for about five minutes.
(5 min)
Now start to bring your awareness back to your surroundings, and gently open your eyes

Control condition

(5 second pause)
Welcome, I will be leading you through this exercise today. To begin, please close your eyes and relax, and let out a long, slow exhale.
(7s)
Relax all the muscles in your body, and take a few long, slow breaths through your nose.
(15s)
While you are sitting quietly, keep your eyes closed and relax as you will listen to information about various relaxation methods, and the science and benefits of these exercises. You will be asked to internally reflect on what you are learning.
(5s)
Many people say that they find listening to music relaxing. To test this, scientists investigated whether listening to classical music would lower people's heart rate, which is an indicator of lowering stress. Each participant listened to Mozart for 20 minutes a day, for 3 days. The scientists compared resting state heart rate before vs. after the three-day intervention to see if it had changed. The results showed that for people over 30 years old, classical music reduced stress. But for people under 30 years old, there didn't seem to be a benefit from listening to classical music. Do you have any ideas why there would be an age effect? Reflect on what might explain this age difference.
(15s)
Do you think this result would change if a different style of music was used? What results would you predict for your favorite style of music?
(15s)
What else, besides heart rate, would you have liked to measure in this study? Do you think listening to classical music would affect anything else?
(20s)
In another study, scientists were interested in whether owning a pet reduces stress. They came up with an idea for a study to test this question: bring participants into the lab and randomly assign them into two groups. One group would receive a pet of their choice to take home with them, and the other group would receive a cute stuffed animal to take home. Then, they would see if the participants who took home a real live pet showed greater stress reduction over time, as compared to those who took home a stuffed animal. Unfortunately, the university at which they worked would not allow the scientists to do this study. Can you imagine some reasons why not?
(20s)
Because the scientists couldn't do the study that they originally planned, they came up with another idea. So, to test whether owning a pet reduces stress, the scientists decided to simply ask whether people who already own a pet report feeling less stressed than people who do not own a pet. To do this, they sent out a survey to 1,000 people in San Diego, and they indeed found that pet owners reported less stress in their lives than people who do not own a pet. The scientists concluded that owning a pet reduces stress. Do you see any problems with this conclusion?
(15s)
Do you think the results would change if they could have done their original study idea? Why or why not?
(15s)
Do you think the type of pet matters for this effect?
(20s)

There have also been many studies showing that exercising at least 3 days a week reduces stress and lowers the chance for cardiovascular diseases. Do you have any personal experience suggesting that this scientific finding is true? Either for yourself or others you know?
(15s)
What do you think of this amount of exercise? Do you think 3 days a week is good for everyone? What amount of exercise is best for you? And what happens when you exercise too little, or too much?
(15s)
How closely related do you think stress and cardiovascular health are? How have you noticed their link, in your life or in others you know?
(20s)
Other studies have shown the importance of sleep for stress reduction, as well as productivity, mood, and immune function. For example, one study showed that students who were sleep-deprived performed worse on a math test than students who had at least 7 hours of sleep. Do you think this result also provides evidence for how sleep deprivation increases stress? Why or why not?
(15s)
What else would you have liked to measure in this study? What other symptoms do you notice in yourself when you don't get enough sleep? How would you measure these symptoms?
(15s)
Do you think people vary in how much sleep they need? How much sleep do you usually need to feel rested?
(20s)
Another scientist was gardening one day and became curious about why he always felt so at peace while working in his garden. He wondered whether it was simply being outside and working with plants, or if it was because he was growing food that he will eat and enjoy later. So, he decided to test this question with a gardening study, where he brought participants into a lab garden to engage with different stages of the growing process. Half the participants planted seeds, and therefore took nothing home with them. The other half of the participants helped with harvest, and each took a small portion home. The scientist measured stress levels before and after gardening, and compared the changes in stress levels of the planters vs. harvesters. The gardening scientist found that both the planters and the harvesters had equal amounts of stress reduction. Therefore, what conclusions do you think this gardening scientist can draw?
(15s)
How generalizable do you think these findings are? Do you think results would change if this study compared different ages or cultures?
(15s)
Can you think of any other explanations for why gardening might be relaxing?
(20s)
Stretching can also be an effective relaxation method. A number of studies have shown that slow stretching or practices like yoga or or tai chi can lower someone's stress response. For example, one study showed that participants who did 30 minutes of stretching had lower cortisol levels than participants who did 30 minutes of reading. Based on this result, how do you think stretching works to relax you? Why do you think stretching is an effective relaxation technique?
(15s)
Do you find this result to be true from your own experience?
(15s)
What other control, besides reading, would you have implemented to compare with stretching?
(20s)

Some people find journaling to be a relaxing or therapeutic exercise. There are many forms and styles of journaling, so one scientist was curious about whether the topic of a journaling session was important for the therapeutic benefits. She brought participants into the lab and had them journal about an inspirational figure, where half of them journaled about a real person and half journaled about a fictional character that they admired. The scientist found that both groups showed equal benefits, and there was no difference between journaling about a fictional or factual person of inspiration. Which do you think is more relaxing, journaling about anything, or thinking about someone who inspires you?
(15s)
What other condition would you add to this study? What else would you ask participants to journal about?
(15s)
Do you think it was easier for participants to come up with a fictional inspirational figure (like Harry Potter) or real life inspirational figure (like Abraham Lincoln)? And for a real life inspirational figure, do you think it's easier to come up with someone from history or someone who is alive today?
(20s)
For about the next 2 minutes, think about whether there are other studies that should be done to test what practices and methods reduce stress. How would you test the effectiveness of a relaxation technique? How would you measure changes in stress? We discussed several biological measures, like heart rate and cortisol levels, as well as subjective measures of stress and inspiration, but what other ways of measuring stress can you think of? Reflect silently, with your eyes still closed, for the last two minutes of this exercise, and think of how you would define and measure changes in stress.
(2 min)
Now start to bring your awareness back to your surroundings, and gently open your eyes.

Appendix B

Original 63 Item Pool

Facet	Trait-FFMQ Referent Item	State-FFMQ Adapted Item
Awareness	5. When I do things, my mind wanders off and I'm easily distracted*	***1. My mind was wandering off and I was easily distracted****
Awareness	8. I don't pay attention to what I'm doing because I'm daydreaming, worrying, or otherwise distracted*	***2. I didn't pay attention to what I was doing because I was daydreaming, worrying, or otherwise distracted****
Awareness	8. I don't pay attention to what I'm doing because I'm daydreaming, worrying, or otherwise distracted*	3. I paid attention and focused on what I was doing
Awareness	13. I am easily distracted*	4. I was easily distracted*
Awareness	13. I am easily distracted*	***5. It was easy to pay attention and focus on what I was doing***
Awareness	18. I find it difficult to stay focused on what's happening in the present*	***6. I found it difficult to stay focused on what was happening in the moment****
Awareness	18. I find it difficult to stay focused on what's happening in the present*	7. I found it easy to stay focused on what was happening in the moment
Awareness	23. It seems I am "running on automatic" without much awareness of what I'm doing*	*8. I was "running on automatic" without being aware of what I was doing**
Awareness	28. I rush through activities without being really attentive to them*	9. I rushed through this experience without being really attentive to it*
Awareness	34. I do jobs or tasks automatically without being aware of what I'm doing*	10. I was behaving automatically without being aware of what I was doing*
Awareness	38. I find myself doing things without paying attention*	11. I was experiencing things without paying attention*
Awareness	38. I find myself doing things without paying attention*	12. I was paying attention to what I was doing
Describing	2. I'm good at finding words to describe my feelings	***13. I would have been good at finding the words to describe my feelings***
Describing	7. I can easily put my beliefs, opinions, and expectations into words	14. I could have easily put my beliefs, opinions, and expectations into words
Describing	12. It's hard for me to find the words to describe what I'm thinking*	***15. It would have been hard for me to find the words to describe what I was thinking****
Describing	16. I have trouble thinking of the right words to express how I feel about things*	16. I would have had trouble thinking of the right words to express how I felt about things*
Describing	22. When I have a sensation in my body, it's difficult for me to describe it because I can't find the right words*	***17. Finding the right words to describe the sensations in my body would have been difficult for me****

Facet	Trait-FFMQ Referent Item	State-FFMQ Adapted Item
Describing	27. Even when I'm feeling terribly upset, I can find a way to put it into words	***18. I would have been able to find a way to put my feelings into words***
Describing	32. My natural tendency is to put my experiences into words	***19. It would have felt natural to put my experience into words***
Describing	37. I can usually describe how I feel at the moment in considerable detail	20. I could have described how I felt in the moment in considerable detail
Nonreactivity	4. I perceive my feelings and emotions without having to react to them	*21. I perceived my feelings and emotions without reacting to them*
Nonreactivity	9. I watch my feelings without getting lost in them	*22. I watched my feelings without getting lost in them*
Nonreactivity	9. I watch my feelings without getting lost in them	23. I got lost in my thoughts or my feelings*
Nonreactivity	19. When I have distressing thoughts or images, I "step back" and am aware of the thought or image without getting taken over by it	24. I "stepped back" and was aware of my thoughts without being taken over by them
Nonreactivity	21. In difficult situations, I can pause without immediately reacting	25. I paused without needing to immediately react to my experience
Nonreactivity	24. When I have distressing thoughts or images, I feel calm soon after	26. My thoughts did not affect how calm I felt
Nonreactivity	24. When I have distressing thoughts or images, I feel calm soon after	*27. I experienced thoughts or images that made me feel less calm**
Nonreactivity	29. When I have distressing thoughts or images, I am able just to notice them without reacting	28. I was able to just notice my thoughts without reacting
Nonreactivity	29. When I have distressing thoughts or images, I am able just to notice them without reacting	*29. I reacted to my thoughts or mental images**
Nonreactivity	33. When I have distressing thoughts or images, I just notice them and let them go	*30. I just noticed my thoughts and let them go*
Nonreactivity	33. When I have distressing thoughts or images, I just notice them and let them go	31. I could not let go of certain thoughts or feelings*
Observing	1. When I'm walking, I deliberately notice the sensations of my body moving	32. I deliberately noticed sensations in my body
Observing	1. When I'm walking, I deliberately notice the sensations of my body moving	***33. I failed to notice sensations in my body****
Observing	6. When I take a shower or bath, I stay alert to the sensations of water on my body	34. I stayed alert to bodily sensations

Facet	Trait-FFMQ Referent Item	State-FFMQ Adapted Item
Observing	6. When I take a shower or bath, I stay alert to the sensations of water on my body	***35. I did not have much awareness of bodily sensations****
Observing	11. I notice how foods and drinks affect my thoughts, bodily sensations, and emotions	***36. I noticed how the experience affected my thoughts, bodily sensations, and emotions***
Observing	11. I notice how foods and drinks affect my thoughts, bodily sensations, and emotions	37. I was unaware of how what I was doing affected my thoughts, bodily sensations, or emotions*
Observing	15. I pay attention to sensations, such as the wind in my hair or sun on my face	***38. I paid attention to sensations***
Observing	15. I pay attention to sensations, such as the wind in my hair or sun on my face	39. I was distracted from paying attention to sensations*
Observing	20. I pay attention to sounds, such as clocks ticking, birds chirping, or cars passing	40. I paid attention to sounds, such as clocks ticking, birds chirping, or cars passing
Observing	20. I pay attention to sounds, such as clocks ticking, birds chirping, or cars passing	41. It was difficult to pay attention to the sounds around me*
Observing	26. I notice the smells and aromas of things	42. I noticed the smells and aromas of things around me
Observing	26. I notice the smells and aromas of things	43. It was difficult to notice the smells and aromas of things around me*
Observing	31. I notice visual elements in art or nature, such as colors, shapes, textures, or patterns of light and shadow	44. I noticed visual elements in the things around me, such as their color, shape, texture, or patterns of light and shadow
Observing	31. I notice visual elements in art or nature, such as colors, shapes, textures, or patterns of light and shadow	45. I was inattentive to visual elements in the things around me or in my thoughts, such as their shape or color*
Observing	36. I pay attention to how my emotions affect my thoughts and behavior	46. I paid attention to how my emotions affected my thoughts or behaviors
Observing	36. I pay attention to how my emotions affect my thoughts and behavior	*47. I failed to notice how my emotions affected my thoughts or behaviors**
Nonjudging	3. I criticize myself for having irrational or inappropriate emotions*	48. I was critical of myself for having irrational or inappropriate emotions*
Nonjudging	3. I criticize myself for having irrational or inappropriate emotions*	***49. Regardless of my thoughts or emotions, I accepted myself***
Nonjudging	10. I tell myself I shouldn't be feeling the way I'm feeling*	50. I should not have been feeling the way I was feeling*
Nonjudging	10. I tell myself I shouldn't be feeling the way I'm feeling*	51. There was no need to change the way I was feeling
Nonjudging	14. I believe some of my thoughts are abnormal or bad and I shouldn't think that way*	***52. Some of my thoughts were abnormal or bad and I shouldn't have thought in that way****

Facet	Trait-FFMQ Referent Item	State-FFMQ Adapted Item
Nonjudging	14. I believe some of my thoughts are abnormal or bad and I shouldn't think that way*	53. My thoughts were normal and there was no need to change the way I was thinking
Nonjudging	17. I make judgments about whether my thoughts are good or bad*	54. I judged whether my thoughts were good or bad*
Nonjudging	17. I make judgments about whether my thoughts are good or bad*	***55. I did not consider whether my thoughts were good or bad***
Nonjudging	25. I tell myself that I shouldn't be thinking the way I'm thinking*	56. I should not have been thinking the way I was thinking*
Nonjudging	25. I tell myself that I shouldn't be thinking the way I'm thinking*	57. There was no need to change the way I was thinking
Nonjudging	30. I think some of my emotions are bad or inappropriate and I shouldn't feel them*	58. Some of my emotions were bad or inappropriate and I shouldn't have felt them*
Nonjudging	30. I think some of my emotions are bad or inappropriate and I shouldn't feel them*	***59. My emotions were normal and there was no need to change the way I was feeling***
Nonjudging	35. When I have distressing thoughts or images, I judge myself as good or bad	60. I judged myself as good or bad depending on what my thoughts were about*
Nonjudging	35. When I have distressing thoughts or images, I judge myself as good or bad depending what the thought or image is about*	61. I accepted myself regardless of what my thoughts were about
Nonjudging	39. I disapprove of myself when I have irrational ideas*	***62. I disapproved of myself for having irrational ideas****
Nonjudging	39. I disapprove of myself when I have irrational ideas*	63. Regardless of my thoughts or emotions, I approved of myself

Note. Items with a * symbol should be reverse coded prior to scoring. Italicized items were retained in the reduced 25 item pool from the EFA analysis. Bolded items were retained in the full-form state-FFMQ.

Appendix C

Measures.

Data Cleaning Measures. The following two items used for data cleaning were displayed on the last page of the survey, with the following header text: “Thank you for your participation in our study! We greatly appreciate it! In order to maintain the validity of our data, please indicate your level of attention and effort while participating in this study. This is still anonymous, and you will get your SONA credit no matter how you answer these two questions, so please be honest. These questions are used to ensure we only include valid data in our analyses. Thank you!”. In both cases participants were excluded if they selected either of the last two response options.

Survey Honesty. “While filling out my responses (before and after the relaxation audio):”, with the following four options: “I read all instructions and questions carefully, and answered honestly to the best of my ability”, “I went through this survey quickly, and may have missed a few instructions or skimmed questions, but I still answered honestly”, “I read most of the instructions and questions, but I clicked some answers without fully reading them”, “I tried to finish this as quickly as possible and did not read most of the questions or instructions, or I did not answer honestly”.

Engagement during audio recording. “While listening to the 20-minute relaxation audio:”, with the following four options: “I listened to all instructions carefully, and participated the entire time to the best of my ability”, “I mostly listened to the instructions, but I did not participate for the entire time”, “I sort of listened to the instructions, but did not participate most of the time”, “I barely listened to the instructions or did not listen at all, and did not try to participate”.

Meditation Experience Items. The following questions were primarily inspired by Baer et al. (2008), Feldman (2010), and Pang & Ruch (2019). First, participants were asked about any previous experience with the question, “Do you have any previous experience with mindfulness or meditation (i.e. any practice that focuses on training attention and awareness with the goal of producing emotional calm, mental clarity, self-awareness, and/or concentration)?” with three options: Yes; Yes, but a while ago; No.

If “Yes” or “Yes, but a while ago” is selected, participants then answered the following questions, with the only difference being the tense of the question: First, “About how frequently do/did you practice mindfulness or meditation?”, with eight options: Several times a day, Once a day, Several times a week, Once a week, Several times a month, Once a month, Several times a year, Once a year or more. Second, “About how much time, on average, do/did you spend practicing mindfulness or meditation per session?”, using a sliding scale of 1 minute units going from 0 - 60+ minutes. Third, “About how long have/had you been practicing mindfulness or meditation?”, which will be presented as a sliding scale with each unit being 1 month, and the range being from 0 - 25+ months. Fourth, “What type(s) of mindfulness or meditation practice do/did you engage in? Please assign a percentage to each type of practice below in terms of how often you engage in it”. This was displayed as a constant sum in Qualtrics, so participants assigned a percentage to each possible type which summed to 100%. The following options were displayed in a randomized order with the exception of “Not sure” and “Other” which were always displayed last: Focused attention meditation (e.g., of the breath or the body); Loving-Kindness or Compassion meditation; Open Monitoring meditation; Mantra or Transcendental meditation; Yoga, Tai Chi, or Qi Gong; Meditation-based religious practices; Not Sure; Other.

If “Yes, but a while ago” was selected, participants were further asked, “About how long ago did you stop practicing mindfulness or meditation?” with the following options: less than 3 months ago, 3-6 months ago, 6-12 months ago, 1-2 years ago, 2-5 years ago, 5+ years ago. For past meditators, most (53.49%) stopped practicing within the past year.

Appendix D

Exploratory Tests Expanding on Discriminant Sensitivity to Mindful States.

Several exploratory tests[5] were run to better understand our null result for the discriminant sensitivity analysis. First, we checked whether the Condition instructions were unclear to any group, which was not found; participants rated the clarity of instructions (scored 1-7) as very clear after both the Meditation (M = 6.20, SD = 0.93) and Control (M = 6.04, SD = 0.99) conditions. Second, because Meditation Status did have a main effect in our models, we checked whether that result was undermined by differences across the Meditation Status groups, but no differences were found; the clarity of instructions did not differ by Meditation Status via a one-way ANOVA, $F(2, 306) = 0.84$, $p = .44$; and there were no significant demographic differences by Meditation Status as assessed with ANOVA and chi-square tests (*p's* > .05). Third, we checked whether there were any major differences in Thought Valence scores between the Conditions, which was not found: mean Thought Valence (scored 0-6) was slightly positive after both the Meditation (M = 3.81, SD = 1.21) and Control (M = 3.61, SD = 1.19).

Next, we explored using the state-MAAS and the SMS total score as dependent variables in the mixed ANOVA to help determine whether the amount of mindfulness induced by our Meditation condition, compared to our Control condition, was strong enough to be detected in extant state mindfulness scales; we found a similar pattern of results with the state-MAAS, and marginally better results with the SMS. When using the state-MAAS score as the dependent variable, the ANOVA did not reveal any statistically significant effects for Condition (Meditation M = 16.71, SD = 6.22; Control M = 17.38, SD = 6.53), $F(1, 614) = 0.03$, $p = .87$), Meditation Status ($F(2, 614) = 0.42$, $p = .66$), or their interaction ($F(2, 614) = 0.21$, $p = .81$). When using the SMS total score as the dependent variable, the ANOVA did not reveal a significant main effect of Condition (Meditation M = 72.98, SD = 12.05; Control M = 70.06, SD = 13.97, $F(1, 614) = 0.16$, $p = .69$, but did reveal a significant main effect for Meditation Status ($F(2, 614) = 7.32$, $p = .001$, $\eta^2 = .02$) and the interaction between Meditation Status and Condition ($F(2, 614) = 3.92$, $p = .02$, , $\eta^2 = .01$). Post-hoc pairwise comparisons using the Tukey method revealed that current meditators exhibited a small but significantly higher total score then non-meditators (mean difference = 5.79, SE = 1.51, $p = .0004$), and past meditators (mean difference = 4.74, SE = 1.96, $p = .04$). No significant difference was found between non-meditators and past meditators ($p = .77$). Several pairwise comparisons were conducted for different combinations of Meditation Status and Condition, but no significant differences were found between these combinations after adjusting for multiple comparisons.

Last, we also explored whether the pre to post benefit from the intervention was different between the Conditions, since only the Meditation was expected to be beneficial to participants. We first established that there were significantly lower scores from pre- to post-intervention for the Meditation condition (State Anxiety: $t(310) = 6.41$, $p < .001$; Cohen's d = 0.36; State Stress $t(310) = 12.25$, $p < .001$; Cohen's d = 0.69). Next, one-sample t-tests revealed that the difference in pre to post difference scores between Conditions was not statistically significant for both State Anxiety ($t(310) = -0.74$, $p = .46$) and State Stress ($t(310) = -1.02$, $p = .31$), suggesting that both Conditions were rated as equally beneficial to participants.

[5] In addition to the exploratory tests described here, pilot data from a different study from our lab (N = 165; in preparation) that included the state-FFMQ at both pre and post single meditation, using the same audio as the current study, was analyzed to determine whether the state-FFMQ detected a pre to post change in state mindfulness. A paired t-test on the total state-FFMQ score showed that pre (M = 14.04) to post (M = 14.36) total scores did significantly increase ($t(126) = 2.60$, $p = .01$), but with a small effect size (Cohen's d = 0.16).

Appendix E

Full-form Results of Further Analyses

Supplemental Table 1: Convergent Validity of the Full-form state-FFMQ

Variable	1	2	3	4	5	6	7	8	9	10	11	12
1. State Awareness												
2. State Describing	.36**											
3. State Nonjudge	.19**	.32**										
4. State Observing	.44**	.39**	.16**									
5. State-MAAS	.74**	.40**	.31**	.47**								
6. SMS Body	.28**	.32**	.10	.60**	.33**							
7. SMS Mind	.38**	.37**	.07	.59**	.34**	.67**						
8. Trait Awareness	.42**	.30**	.24**	.26**	.47**	.20**	.19**					
9. Trait Describing	.17**	.62**	.24**	.23**	.21**	.23**	.29**	.34**				
10. Trait Observing	.22**	.26**	.17**	.43**	.25**	.47**	.39**	.12*	.32**			
11. Trait Nonjudging	.19**	.18**	.49**	.10	.25**	-.04	.06	.42**	.29**	.04		
12. Trait Nonreactivity	.32**	.17**	.29**	.15**	.26**	.11	.18**	.28**	.26**	.36**	.38**	

**p* < .05. ***p* < .01. All correlations had 309 degrees of freedom.

Supplemental Table 2: Test-Retest Reliability of the Full-form state-FFMQ

Measure	Correlation
State-FFMQ Total	0.64
State Awareness	0.53
State Describing	0.60
State Nonjudging	0.56
State Observing	0.42

Discriminant Sensitivity of the full-form state-FFMQ. The analysis revealed a small but significant main effect of Meditation Status on the total state-FFMQ score, $F(2, 614) = 3.38$, $p = .03$, $\eta^2 = .01$. However, there was no significant main effect of Condition, $F(1, 614) = 0.002$, $p = .97$, $\eta^2 < .001$. Additionally, the interaction effect between Meditation Status and Condition was not significant, $F(2, 614) = 0.23$, $p = .80$, $\eta^2 = .001$. Post-hoc pairwise comparisons revealed that current meditators exhibited a small but significantly higher total score than non-meditators (mean difference = 0.17, SE = 0.06, $p = .03$), and trended towards higher total scores than past meditators (mean difference = 0.15, SE = 0.08, p = .17). No significant difference was found between non-meditators and past meditators (p = .97).

Predictive and Incremental Validity of the full-form state-FFMQ.

Supplemental Table 3: Multiple Linear Regression Results Predicting State Anxiety After a Single Meditation Session of the Full-form state-FFMQ

Variable	Predictive Model				Incremental Model			
	B	*SE*	*p*	η2	*B*	*SE*	*p*	η2
(Intercept)	43.03	2.70	<.001		42.04	3.11	<.001	
Pre-Score	0.35	0.03	<.001	.34 [.26, .42]	0.36	0.03	<.001	.35 [.27, .42]
Awareness	-1.44	0.34	<.001	.06 [.02, .11]	-1.22	0.34	<.001	.04 [.01, .09]
Describing	-0.25	0.37	.51	.001 [.00, .02]	-0.34	0.37	.36	.003 [.00, .03]
Nonjudging	-1.94	0.41	<.001	.07 [.02, .13]	-1.87	0.41	<.001	.06 [.02, .12]
Observing	-1.53	0.47	.001	.03 [.01, .08]	-1.52	0.47	.001	.03 [.01, .08]
Trait-FFMQ Total					0.03	0.02	.17	.006 [.00, .04]
Thought Valence					-0.81	0.23	<.001	.04 [.01, .09]
R2				0.57				0.59

Note. N = 311. The F-statistic is significant at $p < .001$ in all models. *B* represents unstandardized regression weights. η2 represents partial eta-squared. Numbers in brackets represent 95% confidence intervals. Collinearity as assessed with VIF were well within acceptable limits for each model. Predictors were entered into analysis simultaneously.

Supplemental Table 4: Multiple Linear Regression Results Predicting State Stress After a Single Meditation Session of the Full-form state-FFMQ

Variable	Predictive Model				Incremental Model			
	B	*SE*	*p*	η2	*B*	*SE*	*p*	η2
(Intercept)	10.31	0.90	<.001		10.70	1.01	<.001	
Pre-Score	0.36	0.04	<.001	.23 [.15, .30]	0.36	0.04	<.001	.23 [.16, .31]
Awareness	-0.60	0.14	<.001	.06 [.02, .12]	-0.48	0.14	<.001	.04 [.01, .00]
Describing	-0.10	0.15	.50	.002 [.00, .02]	-0.08	0.16	.60	.0009 [.00, .02]
Nonjudging	-0.63	0.16	<.001	.05 [.01, .10]	-0.54	0.17	.001	.03 [.01, .08]
Observing	-0.74	0.19	<.001	.05 [.01, .10]	-0.71	0.18	<.001	.05 [.01, .10]
Trait-FFMQ Total					-0.002	0.01	.97	.000004 [.00, .00]
Thought Valence					-0.32	0.09	<.001	.04 [.01, .09]
R2				0.46				0.47

Note. N = 311. The F-statistic is significant at $p < .001$ in all models. *B* represents unstandardized regression weights. η2 represents partial eta-squared. Numbers in brackets represent 95% confidence intervals. Collinearity as assessed with VIF were well within acceptable limits for each model. Predictors were entered into analysis simultaneously.

Chapter 2 Multiple Facets of Daily Mindfulness Uniquely Predict State Affect

Abstract

The assessment of state mindfulness often involves detecting mindful experiences during formal practices like meditation. However, state mindfulness is also experienced moment by moment throughout typical daily experiences. The current research employed the Day Reconstruction Method (DRM) to assess whether different components of state mindfulness, captured via reflections of experiences from the previous day, uniquely predict state affect, in particular, state happiness. A secondary aim was to determine whether there was a long-term benefit to completing the DRM, compared with two controls. Undergraduate students completed measures of trait mindfulness and happiness on days 1, 4 and 11, with the goal of investigating whether participating in the DRM (administered on days 2 and 3) caused improvements on these measures. For each recalled experience, participants reported on state mindfulness, state affect, as well as other covariates likely to predict state affect. State mindfulness was assessed with the newly created State Four Facet Mindfulness Questionnaire, adapted from the commonly used trait Five Facet Mindfulness Questionnaire. Multilevel modeling revealed that daily experiences of state Awareness and Nonjudgement, and to a lesser degree Observing, uniquely predicted state affect. Describing was not uniquely predictive. Various analyses on the current data also provided further validation of the state-FFMQ. Analysis of Variance (ANOVA) tests revealed there was no long-term benefit to completing any condition per se. Different components of daily mindfulness uniquely contribute to state affect, even when accounting for relevant covariates. These effects can be accurately assessed by repeatedly measuring the state-FFMQ within the DRM.

Introduction

Are people happier when they are more mindful? Research has generally explored this question within one of two contexts. The first context is "formal mindfulness", which are structured practices like meditation that are specifically designed to cultivate one or more aspects of mindfulness. Formal mindfulness is primarily studied through experiments that purposely induce mindful states, or through interventions such as Mindfulness-Based Stress Reduction (Kabat-Zinn, 1982), and has the benefit of allowing researchers to control and standardize many aspects of the experimental design. The second context is "daily mindfulness", which are informal, continuous, and spontaneous instances of mindful states that fluctuate in intensity throughout the day. Though less studied than formal mindfulness, daily mindfulness aligns with the highly referenced conceptualization of mindfulness as an inherent and universal human capacity (Brown & Ryan, 2004), experienced moment by moment to varying degrees (Kabat-Zinn, 2003). Daily mindfulness is primarily studied either through correlational designs that assess mindfulness as a *trait* (i.e. a retrospective self-assessment of how typical it is to experience mindfulness most of the time); or through a more rigorous and ecologically valid quasi-experimental design, which assesses specific instances of mindfulness as a *state* (i.e. a self-assessment of one's experience of mindfulness in reference to the current moment, or in relatively close temporal proximity to the current moment). Thus, to determine whether people are happier when they are more mindful in a methodologically rigorous and ecologically valid way, state measures of daily mindfulness are preferrable.

One of the most frequently used methods that assesses daily mindfulness as a state is the quasi-experimental intensive longitudinal design (ILD), where participants rate aspects of their daily experience repeatedly over time. The rich datasets that result from ILDs can examine short-term, within-person, processes that best represent the ebb and flow of daily mindfulness (see

Schneider et al., 2020). Although the advantages of ILDs are well-established, the choice of which ILD method to use is more ambiguous, with two[6] main methodologies available (see more detailed overviews of ILD methods in Bamberger, 2016; Bolger & Laurenceau, 2013; Schneider et al., 2020; Schneider & Stone, 2016): the Experience Sampling Method (ESM; Stone & Shiffman, 1994) and the Day Reconstruction Method (DRM, Kahneman et al. 2004).

In ESM (ESM; Stone & Shiffman, 1994), participants are prompted throughout the day (typically for several consecutive days) to provide data about their behaviors and experiences as they are happening in real time or in close proximity to real time (e.g., in the moment just prior to being prompted). Because ESM captures the experience in situ, it can be relatively easy for participants to complete, and can be coupled with other real-time ambulatory measures (e.g., heart rate variability), it is typically considered the gold standard among ILD methods. The question of whether people are happier when they are more mindful has been amply explored with ESM designs (e.g., Raugh et al., 2023), with one systematic review identifying 22 articles that used ESM to investigate the effects of daily mindfulness or mindfulness training on mental health outcomes (Enkema et al., 2020). Though the review did not distinguish between daily contexts (i.e., studies that prompt participants throughout their day) and formal contexts (e.g., studies that prompt participants during a mindfulness-based intervention), it overall found consistent positive associations between mindfulness and mental health outcomes (including state affect) with relatively large effect sizes. Although these effects are promising, one gap in this literature is the inconsistency in how state mindfulness has been operationalized. In a similar

---

[6] Note that a third method, Daily End-of-Day diary (EOD; also called the daily diary method), are not discussed since the mindfulness literature has largely moved away from this less rigorous ILD method in recent years. In EOD designs, participants respond to questions about one's behaviors and experiences in reference to the entire day (e.g., how stressful was today), over the course of multiple days. EOD is the least valid ILD method because it relies on recall alone rather than recall after episodic reinstantiation (DRM) or momentary recall (ESM), thereby making it the most prone to bias.

vein, state mindfulness is likely to be multidimensional (as is argued to the be the case for trait mindfulness; see Bergomi et al. (2013), and it is therefore of interest to know *which* aspects of state mindfulness are related to state affect. Unfortunately, most ESM studies have not investigated the relationship of *different* state mindfulness facets on state affect, which may be attributed to the lack of a validated multidimensional state mindfulness measure that is short enough to be employed in ILDs without overburdening participants. Therefore, one of the main goals of the current study was to ask whether different facets of state mindfulness uniquely predict state affect, in particular, happiness, and for this we chose to use the DRM, rather than the ESM, paradigm based on some shortcomings of ESM[7].

Kahneman et al. (2004) designed the DRM as a means of reproducing the information that would be collected through ESM, but without the shortcomings described in footnote 2. Based on techniques grounded in cognitive science (see commentary by Diener & Tay, 2014; Krueger et al., 2004; Ludwigs et al., 2019; Schwarz et al., 2009), participants first systematically reconstruct the previous day into specific, sequential, single episodes (calling upon episodic memory); then they report on their behaviors and experiences during each individual episode. Early studies validated the DRM by showing that changes in affect collected over the course of

[7] Although individual prompts are typically short, ESM studies are often *overall* described as time-consuming (see Kahneman et al., 2004) and burdensome to participants, and therefore has high potential for attrition. Participants can also be burdened by the intrusiveness of needing to carry and attend to a device that repeatedly interrupts their day at unpredictable times to complete the same prompted survey (see Hudson et al., 2020). Such interruptions may lead to assessments being biased due to reactions from the notification per se (e.g., feeling annoyed at the sound of a notification going off at an inconvenient time), and over time ESM reporting may lead to participants paying more attention to their moods and emotions throughout the day, thereby limiting the ecological validity of results (see Diener & Tay, 2014). There are also circumstances where real-time data collection is not feasible, such as among individuals without access to a smartphone/wearable tech, or with certain occupations (e.g., truck driving) or disabilities. Furthermore, because of the randomness by which prompts are delivered to participants, ESM has been criticized for only showing excerpts of daily life depending on when the prompt is responded to, rather than depicting the whole day, thus prohibiting precise time use information (see Kahneman et al., 2004). Last and perhaps most compelling to researchers, it is highly expensive and resource-intensive to conduct an ESM study, particularly if one needs to provide smartphones or wearable tech to participants in addition to purchasing and learning to use the ESM software itself.

one day closely corresponded to data collected from separate studies, and separate samples, that used the ESM methodology (Kahneman et al., 2004; Stone et al., 2006). Since then, a handful of studies have more directly compared reports collected from different ILD methods from the same participants over the same time. Results have generally shown that aggregate measures of affect are in high agreement between ESM and DRM methods, while within-person differences in affect are in somewhat lower agreement method the two methods (Bylsma et al., 2011; Dockray et al., 2010; Kim et al., 2013; Lucas et al., 2021; Schneider et al., 2020). Notably, these results differ somewhat by which aspect of state affect was assessed; for example, Dockray et al. (2010) found that happiness was indistinguishable whether measured with ESM or DRM. Though more research is needed, the findings from these studies suggest that DRM and ESM methods produce relatively comparable results.

Beyond collecting similar datasets as ESM, the DRM has several unique practical benefits as a day-recall ILD method. First, the DRM has a low response burden with minimal invasiveness since it is completed in a single sitting, usually over the internet, at a time that is chosen to be convenient and without potential for interruption by the participant. This autonomy prevents most of the burden involved in responding to automated prompts throughout the day as with ESM (see a similar point made by Oerlemans & Bakker, 2013). Second, the DRM is free to administer and does not require additional software to complete. For these reasons, the DRM is usable in national surveys (e.g., Hudson et al., 2017), a context that precludes use of ESM given its scale (see Kahneman et al., 2004). Third, the DRM has more complete coverage of a typical day since it assesses episodes from the entire day from waking to sleeping in chronological order (noting ESM can also do this with a very dense sampling scheme, but that would drive up participant burden). DRM also includes precise information about the duration of each episode, which can be used in various research applications such as duration weighted analysis. Last and most relevant to the current study, the measurement of daily mindfulness may *best* be captured by episodes rather than specific in situ moments. Commenting on the development of the Multidimensional State Mindfulness Questionnaire, which employed ESM, Blanke and Brose (2022) highlighted that in participant feedback, mindful experiences were reported as being better and more preferentially assessed during time frames, rather than pinpointed to one specific moment. The authors speculated that a mindful state likely spans a longer time frame than moments (as captured by ESM), such as hours (which is best captured by DRM); however, there is a dearth of empirical research assessing the duration of mindful states. Although gradually becoming a more well-validated ILD method (see Ludwigs et al., 2019), the DRM has garnered little attention in the literature, highlighting the need for further research using the DRM.

For all these reasons, the current study employed DRM to ask whether people are happier in daily life moments when they feel more mindful, and moreover whether different components of state mindfulness differentially predict happiness. To the best of our knowledge, this is the first study measuring daily mindfulness in a DRM design. For this study, we employed a recently developed state mindfulness scale; the Four Facet Mindfulness Questionnaire ("state-FFMQ"; Raynes & Dobkins, under review), which was adapted from the most commonly used multidimensional trait measures of mindfulness: the Five Facet Mindfulness Questionnaire ("trait-FFMQ"; Baer et al., 2006). This new state mindfulness scale was validated using EFA/CFA, and shows good construct, convergent, predictive, and incremental validity. It is also brief enough to be readily used in the DRM or any ILD, and the items were created to be applicable to diverse situations (e.g., formal vs. daily mindfulness) and amongst a general population. Both the state- and trait- FFMQ include the following four[8] facets: *Awareness* is the attention one pays to the present moment, as opposed to focusing attention elsewhere or behaving automatically. *Describing* refers to the ability to express one's experiences in words. *Nonjudging* of inner experience is the acceptance of one's thoughts and emotions without evaluation. Last, *Observing* specifies attending to or noticing both internal and external experiences, such as thoughts, emotions, bodily sensations, smells and sounds. In the development article of the state-FFMQ (Raynes & Dobkins, under review), state affect following a single meditation session was strongly and equally predicted by Awareness and Nonjudgement, followed closely by Observing, and with Describing not being a significant predictor. These results were resilient to the inclusion of several relevant covariates. Thus, the development article captured the unique predictive effects of each state-FFMQ facet in the context of formal mindfulness. The current study asks this same question but in the context of daily mindfulness, with the expectation of finding the same pattern of results.

Because the newly created state-FFMQ has only been validated in its development article, more studies are required to provided further validation. This was the second goal of the current study, i.e., to test the predictive, incremental, convergent and construct validity of the state-FFMQ in the context of daily mindfulness.

The third goal of the current study was to test if there are long-term benefits on trait mindfulness and happiness of participating in the DRM, which was inspired from two sources. First, Bergomi et al. (2013) proposed in a review paper that the act of responding to mindfulness questionnaires may itself aid in the development of trait mindfulness. Second, we hypothesized

[8] The trait-FFMQ has a fifth Nonreactivity to inner experience facet, which is the ability to allow thoughts and emotions to come and go without becoming attached or carried away with them. This facet was eliminated in the development of the state-FFMQ.

that the practice of reconstructing the details of one's day through the DRM may itself be an "intervention" protocol by focusing attention on internal experiences during daily experiences, similar to daily journaling, and therefore may be psychologically beneficial. As such, we predicted that participating in the DRM protocol per se, compared with two control protocols, would lead to small but significant increases in trait mindfulness and trait happiness. To the best of our knowledge, this is the first study to investigate whether participation in an ILD method per se affects trait outcomes relative to controls.

Method

Participants. Participants were undergraduate students recruited in 2023 - 2024 through the UCSD SONA 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. All participants gave their informed consent before participating and were compensated with course credit.

Sample size was a priori determined based on pilot data collected in our lab, which showed significant effects with a sample size of 104 (after cleaning). It was therefore our goal to obtain useable data from 105 participants in the current study, for each of the three conditions (see Procedures, below). We chose this method for determining sample size after Blanke et al. (2018) rather than using a formal power analysis, because the latter is complex and controversial for multi-level models (see Aguilar-Raab et al., 2021; Mathieu et al., 2012). To obtain usable data for the three conditions, we aimed to collect data from 371 participants, so that after an expected loss of 15% (from attrition and after data cleaning), we would end up with at least 105 per condition. The collected sample consisted of 416 participants.

The following five exclusion criteria (as outlined in our pre-registration) were applied to the total collected sample. *First,* 19 participants were excluded for failing to complete the entirety of the study. *Second*, nine participants were excluded for failing to complete the study within +/- 3 standard deviations of the median study duration (differentiated by condition). *Third*, two participants were excluded for failing to correctly respond to at least one out of two attention check questions dispersed throughout the study. *Fourth*, 10 participants were excluded for admitting (at the end of the study) to not answering the survey questions honestly and attentively (see wording in Raynes & Dobkins, under review). *Fifth*, similar to Ludwigs et al. (2019), six participants were excluded for failing to list more than one episode in their day (see Procedure, below). In sum, a total of 46 participants were excluded for not passing these criteria. While we acknowledge that our exclusion criteria are strict and therefore limits 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 total final sample thus consisted of 370 participants.

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. This was a quasi-experimental study design. Due to logistical constraints of the SONA system, true random assignment of each participant to one of the three conditions was not possible. Therefore, eligible participants signed up for one of three available SONA studies (and were unable to sign up for more than one study), which differentiated their condition. All studies were listed as a five-day study, and were worth the same amount of course credit, had the same study abstract, and had as similar a study description as possible. In an effort to reduce participant attrition, automated email reminders from Qualtrics were sent to participants to complete each part and credit was assigned only after completing all five days.

For each of the three groups of participants, the study was self-administered over the course of 11 days. We refer to days 1, 4 and 11 as “Pre-Intervention”, “Post-Intervention” and “Follow-Up”, respectively, noting that the protocol for these sections was identical across the three participant groups. We refer to days 2 and 3 as the “Intervention”, noting that here the protocol *differed* across the three participant groups. This design allowed us to ask our three main questions. First, data from days 2 and 3 collected from participants in the ”DRM condition” (see below), allowed us to ask whether the different components of state mindfulness uniquely predict state affect, specifically happiness, in daily life. Second, the data from days 1, 2 and 3 from participants in the ”DRM condition” allowed us to conduct further validation analyses of the newly created state-FFMQ. Third, data from days 1, 4 and 11 allowed us to ask if there are long-term benefits of participating in the DRM.

Pre-Intervention (day 1): The order of events was as follows. Participants first filled out a trait measure of mindfulness and a trait measure of happiness, which were randomized in order. Next, they filled out standard questions about demographics, and last, questions about previous meditation experience.

Intervention (days 2-3): As mentioned above, the intervention protocol differed across the three groups of participants. Participants in the “DRM condition” completed an electronic diary, that is, an online version of the DRM (Kahneman et al., 2004). Note that although the DRM is typically administered over a single day, the current study had participants complete the DRM for two consecutive days, which was inspired by past studies (e.g., Dockray et al., 2010; Ludwigs et al., 2019) utilizing the DRM on successive days in order to get a larger and more representative dataset of daily life experiences. We also reasoned that extending the intervention protocol duration might increase our chance of seeing a long-term benefit of participating in the DRM.

Each day the DRM was completed, participants were first asked what time they woke up and went to sleep on the previous day. Next, they were asked to “think of yesterday as a series of scenes in a movie” and to divide the day into separate “episodes". It was explained to them that many people define episodes that last between 15min – 2hrs, yet they were nonetheless encouraged to define episodes, in whatever time bins, as made most sense to them. Beginning with the time they woke up and ending with the time they went to sleep, participants used an open-ended text entry to provide a label for each episode to describe what they did during that time. The open-ended text entries were not included in data analysis and were only for the benefit of the participant, which was made clear in the instructions. In addition to providing a label, for each episode, participants also reported 1) Time: the approximate state and end time (dropdown selection in 5-minute increments), which was used to help the participant recall the episode into episodic memory, 2) Remember: “How much of this episode do you remember?” (rated on a 5-point Likert scale with five labels: None of it, Very little of it, Some of it, Most of it, All of it), which was used as an integrity check to ensure participants recalled the episode they were reporting on in sufficient detail as to be valid (see below), and the 3) Activity Valence (which was used as a covariate, see Data Analysis). Participants could describe up to thirty episodes. If a participant was awake for over 24 hours, they were asked to enter episodes for the first 24 hours they were awake. This portion of listing out the episodes is referred to as the “reinstantiation task”.

After the reinstantiation, they then completed what we refer to as the “mindfulness reporting task”, performed separately for each of their listed episodes (with each episode presented on a separate page). To improve the integrity of the data, only episodes that were sufficiently remembered by the participant (which we defined as episodes in which they selected one of the top three choices in the Remember question, see above) were included for this task[9]. For each sufficiently remembered episode, participants answered a single question about state affect. This was followed by questions pertaining to state mindfulness. As a last question, they were asked to report on the Activity Type the episode could be categorized as.

The two other groups of participants were placed in one of two control conditions, the "Close control" and the "Far control" conditions. In the Close Control condition, on both days 2 and 3, participants *only* completed the "reinstantiation task" described above, and not the "mindfulness reporting task". This control group (sometimes referred to as an "active" control group) allowed us to test whether any long-term benefits of participating in the DRM was attributable to the elaborative reflecting of one's internal experiences during the day's episodes (unique to the DRM condition), as opposed to the cognitive and more factual recollection of one's activities during the episodes (true for both the DRM and the Close Control condition). In the Far Control condition, participants did not complete any task on days 2 and 3. This "passive" control group allowed us to test whether any benefit of completing the DRM was not simply due to chance, time passing, or the experience of repeatedly answering the trait measures.

Post-Intervention (day 4): Participants in all three groups answered the same two trait measures from the Pre-Intervention (day 1) in a randomized order, then an item about the typicality of the prior two days.

[9] After completing the reinstantiation task, participants in the DRM group were instructed that they would be asked questions about a random drawing of listed episodes. In reality, all episodes that passed this integrity check were included. Though possibly annoying, this was done in an attempt to avoid participants from figuring out that they could purposely "fail" this integrity check and speed through the survey on day 3.

Follow-up (day 11): One week later, participants in all three groups answered the same two trait measures from Pre- and Post- Intervention (days 1 and 4), in a randomized order. Last, they answered the survey honesty and attention item used for data cleaning. The Far Control condition then additionally completed a handful of unrelated and unanalyzed surveys so that the total time commitment (and thus the amount of course credit) would be the same across the three groups of participants.

Measures.

Trait Measures. These measures were asked at Pre-Intervention, Post-Intervention, and Follow-up (days 1, 4 and 11), for all three groups.

Trait mindfulness. The 15-item Five Facet Mindfulness Questionnaires (trait-FFMQ; Baer et al., 2012; Gu et al., 2016), which captures the following five dimensions of trait mindfulness on a Likert scale ranging from 1-5: Observing, Describing, Acting with Awareness, Nonjudging, and Nonreactivity. Responses to each facet and the total score are averaged, with higher scores reflecting greater trait mindfulness. The scale showed acceptable internal consistency with Cronbach's alpha coefficients ranging from $\alpha = 0.77 - 0.84$ for the total across the three timepoints in the present study.

Trait Happiness. The 4-item Subjective Happiness Scale (SHS; Lyubomirsky & Lepper, 1999). This measure assesses subjective perceptions of global happiness on a Likert scale ranging from 1-7. Responses are averaged to provide a single total score, with higher scores reflecting greater overall happiness. The scale showed acceptable internal consistency with Cronbach's alpha coefficients ranging from $\alpha = .85 - .87$ across the three timepoints in the present study.

Reinstantiation Task Measures. For the reinstantiation task (which both DRM and Close Control participants completed), for each episode, participants were asked to label the episode, select the start and end time, and select how well they remembered the episode (see Procedure, above). They were also asked:

Activity Valence. 'How would most people rate this activity (regardless of how it was for you)?", rated on a seven-point Likert scale with seven labels ranging from Extremely unpleasant to Extremely pleasant. This item was inspired by a previous study that controlled for "daily event negativity" on a memory recall task (Colombo et al., 2024), which we expanded to include positivity as well. Because it is impossible to objectively measure how pleasant or unpleasant a given activity is for all people, we tried to ask the participant to think about the activity in the third person and as "most people" would rate it, which although is imperfect, can be considered an approximate way to differentiate the activity per se from one's feelings during the activity.

Mindfulness Reporting Task Measures. For each episode, participants completed the following three measures in the following order. The header text read, "During Episode [number], which lasted from [start time] to [end time]:".

State Affect. This item asked: "Please indicate how happy you were feeling in that moment during the episode", which was answered on a 5-point bipolar sliding scale with a resolution of 0.1 and with three labels: Very Unhappy (-2), Neutral (0), and Very Happy (2) (with numbers hidden from participants). The use of this single item was inspired from a recent DRM study using a similarly worded single item for state affect as the dependent variable (Henwood et al., 2022). Further support for the use of a single item for this construct comes from two sources: another DRM study demonstrating that a single-item happiness measure was strongly correlated with a multi-item happiness measure, suggesting that the two are interchangeable (Knabe et al., 2010); and the commonality of single-item measures of positive affect in ILD studies (e.g., Ludwigs et al., 2019) as a means of reducing participant burden by shortening the repeated survey.

State Mindfulness. We used the 12-item short-form[10] of the State Four Facet Mindfulness Questionnaire (state-FFMQ; Raynes & Dobkins, under review), which is a state adaptation of the Five Factor Mindfulness Questionnaire (trait-FFMQ; Baer et al., 2006), and employs four of the five facets (Awareness; Describing; Nonjudging; Observing; and omitting Nonreactivity; see Introduction) with three items per facet. Participants were asked to "Please rate each of the following statements with the number that best describes your own opinion of what was true for you in the moment during the episode". All items were answered on a sliding scale from 1-5 with a resolution of 0.1 and with three labels: Not at all (1), Moderately (3), Completely/Entirely (5). To reduce participant burden and be consistent with other ILD studies, items were arranged in a nonrandom order that alternated among the four facets.

Activity Type. "Whether in-person or online, what type of activity were you primarily doing? (select one)", with six options adapted from an empirical study that used the DRM (Oerlemans et al., 2011) and used in a recent ESM study (Gross et al., under review): 1) Social activity, but NOT engaged in conversation (e.g., being around family, friends, or peers but not speaking; listening to others at a party or group outing); 2) Physical activity (e.g., exercising, sports, walking, bicycling, hiking); 3) Restful activity (e.g., eating, resting, taking a nap, doing nothing, reading for fun, watching TV or videos, browsing the internet or social media); 4) Household activity (e.g., preparing meals, grocery shopping, household finances, cleaning or

[10] All 18 items of the full-form state-FFMQ were asked in the DRM, but the current study only uses data from the 12-item short-form since the short-form is more likely to be used in future ILD studies. There were no substantial differences in results when using either version of the state-FFMQ.

other chores); 5) Cognitive activity (e.g., studying, homework, attending lecture, learning something new, puzzle solving; 6) Other activity (text entry).

Descriptive Measures. The following exploratory measures were collected in all three groups of participants.

Meditation Status. Asked at Pre-Intervention, day 1, these questions were primarily inspired by Baer et al. (2008), Feldman (2010), and Pang & Ruch (2019). Participants were first asked, "Do you have any previous experience with mindfulness or meditation (i.e. any practice that focuses on training attention and awareness with the goal of producing emotional calm, mental clarity, self-awareness, and/or concentration)?" with three options: Yes; Yes, but a while ago; No. If "Yes" or "Yes, but a while ago" was selected, participants then answer the following question, with the only difference being the tense of the question: "About how frequently do/did you practice mindfulness or meditation?", with eight options: Several times a day, Once a day, Several times a week, Once a week, Several times a month, Once a month, Several times a year, Once a year or more. 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, Raynes & Dobkins, under review), 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". 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 explored in future studies. We asked this for descriptive purposes and to be used in an exploratory test of construct validity via the discriminant sensitivity of the state-FFMQ to detecting mindful states.

Typicality of Days. Asked at Post-Intervention, day 4: "We'd like to know if any major event affected you in the past few days (e.g., having stressful interviews, new health issues, or other major changes in your daily routine, etc.)?" with the following three response options: No, the past two days were fairly typical; Yes, there was a major event and it was quite upsetting to me; Yes, there was a major event and it was quite wonderful to me. We asked this for descriptive purposes.

Data analysis.

General. 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 the following effect sizes rather than statistical significance since the latter is often misleading: Pearson r values for bivariate correlations, with the rule of thumb that absolute values of .10 - .30 are weak effects, .30 - .50 are medium effects, and .50 and over are 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 ($\eta 2$) for analysis of variance (ANOVA) tests, 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).

Several analyses employed multilevel models (MLM) since our study design has a natural two-level structure, where prompts collected over time are nested within individuals. All analyses were computed using R (Version 4.2.2; R Core Team, 2022) and the R-package lme4 (v1.1-27.1; Bates et al., 2015) with a maximum likelihood method of estimation and using Type III sum of squares. Prior to analysis, all continuous level 1 variables were person-mean centered, sometimes referred to as “centering-within-clustering”, which reveals within-person effects while eliminating level 2 (i.e., between-person) effects in a multilevel model (Enders & Tofighi, 2007; Nezlek, 2012). For analyses involving level 2 effects, overall mean scores across episodes were used rather than duration-weighted mean scores across episodes as suggested by Kahneman et al. (2004). This is because the validity of weighting by episode duration has been questioned by other researchers (e.g., Diener & Tay, 2014; Henwood et al., 2022), and since there are no established best practices for how to estimate overall mean scores from repeated momentary ratings that are superior to simple mean ratings across episodes (see Hudson et al., 2020). For consistency, all MLMs used fixed slopes, with participant ID entered as a random intercept effect. Following the methodology used by Blanke et al. (2018), effect sizes were calculated via likelihood-ratio based pseudo-R^2 estimates, which approximates the unique variance accounted for by each predictor variable in a MLM by sequentially removing one predictor variable at a time and comparing the R^2 statistics of the nested models (i.e. the full model versus a model with one variable removed, noting that when there was only one fixed effect predictor variable, the comparison was made with the null model). This statistic helps reveal the relative importance of each predictor variable in a model. The assumption of dependency was confirmed in the null model, with the ICC revealing that 16.45% of the variance in state affect was due to between-person variance. No model presented violations of these assumptions: linearity, homoscedasticity, multicollinearity, or normality of residuals, predictor or dependent variables.

Do Different Components of State Mindfulness Uniquely Predict State Affect? To test this, we measured the unique predictive effects of each of the four state mindfulness facets (entered simultaneously as fixed effects) on state affect in a MLM, with the expectation of detecting varying effect sizes across the facets. We expected a positive relationship between the state-FFMQ and state affect.

The resilience of these results was tested by assessing whether the state-FFMQ facet scores still uniquely predicted variance in state affect after accounting for several relevant fixed effect covariates. The first set of covariates was 1) Activity Type, i.e., the type of activity one was engaged in and 2) Activity Valence, i.e., the valence of that activity, which were included in a "robustness" model (see below) since previous ESM studies have shown that activity type (e.g., Gross et al., under review; Killingsworth & Gilbert, 2010) and activity valence (e.g., Colombo et al., 2024) demonstrate significant associations with measures of state affect (the dependent measure in the current study). These covariates were expected to show strong main effects in the current study since state affect should be closely related to what one is doing in that moment. With that in mind, the purpose of including these covariates was to minimize unaccounted for variance in state affect scores, so that we could more clearly measure the direct effects of the state mindfulness facets on state affect. In our sister paper (Raynes & Dobkins, under review), the related construct of *thought* valence during a meditation session showed a strong main effect in predicting state affect, but did not impact the strength of the relationship between state mindfulness facets and state affect.

The second covariate was Trait Mindfulness[11] (trait-FFMQ, Baer et al. 2006, obtained on day 1), which was included in a "incremental" model (see below) as a means of disentangling the effects of state versus trait aspects of mindfulness, which is specifically relevant to "Construct

[11] Although there was a between-person effect of the average state-FFMQ score over the two days (i.e. the level 2 effect) that we could have used, we opted instead to operationalize trait mindfulness using the day 1 trait-FFMQ score, which was assumed to capture trait mindfulness in a more valid way.

Validity" (below).

Given that any of these variables (trait-FFMQ, Activity Type, Activity Valence) may also share variance with state mindfulness (the predictor variable in the current study), their inclusion allows us to pull out the *unique* contribution of state mindfulness to state affect. To confirm these variables were suitable to be included as covariates, we first examined bivariate associations between state affect and the three covariates. Due to the repeated testing nature of the study, we could not rely on bivariate correlations. Instead, we employed three MLM's that included one potential covariate (trait-FFMQ total score, Activity Valence, or Activity Type) and the dependent variable (state affect). We separately tested whether a potential covariate interacted with the state-FFMQ total score in predicting state affect; if it did interact, that term would not be used in the model as a covariate. Given that these variables were confirmed as suitable covariates, we ran an incremental model (which only included the trait-FFMQ) and a robust model (with Activity Valence and Activity Type, entered simultaneously).

Further Validation of the Newly Created State-FFMQ. Because the State-FFMQ is a new scale, we took advantage of the data collected in the current study to further validate the scale (all of which was pre-registered). Note that for these analyses we used data only from participants in the DRM condition, as this was the only condition that used the newly created state-FFMQ.

Predictive validity. If we find evidence that several components of state mindfulness uniquely predicted state affect, this would provide evidence of the predictive validity of the state-FFMQ, as was observed in the development of the state-FFMQ (Raynes & Dobkins, under review).

Construct validity. Construct validity of the state-FFMQ was examined in three ways. First, the multidimensionality of the state-FFMQ was tested. Specifically, we asked whether the test of predictive validity (above) had a superior fit statistics to an alternative one-facet model (i.e. an MLM of the total state-FFMQ score). We then quantified the improvement in model fit by computing the differences of respective AIC values and subjecting the difference to a chi-squared test. This would dismiss the idea that a unidimensional state mindfulness value is a better predictor than a multidimensional scale in our model, despite the additional parameters that the multidimensional model would have.

Second, we tested whether the state-FFMQ was behaving in a state- rather than trait-like manner. If, in the incremental model, the state-FFMQ predicts state affect over and beyond that predicted by the trait-FFMQ, this would provide evidence that the state-FFMQ is not masquerading as a trait measure. This was also assessed by testing whether each component of the state-FFMQ sufficiently varied within a person during instances of daily life. This was assessed with the intraclass correlation coefficient (ICC) for each facet in a multilevel model, which revealed the percentage of variance in a variable that is due to within versus between person variance. A null model was run for each facet, where the facet was the dependent variable, and no predictor variables were added. The majority of variation in each facet was expected to be due to within-person variability, and not between-person variability. Though no specific ICC cutoff exists for this purpose, relatively lower ICC values (~ < .50) suggest that the variable is more indicative of a state than a trait.

Third, though not pre-registered, we explored the discriminant sensitivity of the state-FFMQ to detecting mindful states. Inspired by Burzler and Tran (2022), the mean state-FFMQ total score was expected to be greater as meditation experience increases (i.e. those with current or past meditation experience should, on average, be more mindful in daily life than non-meditators) as assessed by an ANOVA on the level 2 state-FFMQ total score. Note however that we expected the majority of our sample to be classified as non-meditators given our convenience sampling of undergraduate students.

Convergent validity. Last, we were also able to perform a test of the convergent validity of the state-FFMQ by asking whether each state-FFMQ facet aligned well with the same facet from the trait-FFMQ facet. For this, we calculated bivariate correlations between the level 2 state-FFMQ facets (i.e., the between-person average across episodes) and the trait-FFMQ facets (obtained from participants on day 1), with the expected result of a moderate to strong and positive correlation amongst each aligned facet (e.g., between mean state Awareness vs. trait Awareness). However, the effect size of the expected convergences was not expected to be too strong, as previous literature suggests that global self-reports and aggregated state averages assess somewhat differing constructs (Blanke et al., 2018; Hudson et al., 2020; Robinson & Clore, 2002b, 2002a).

We also separately explored the convergent validity of our in-house state affect measure using this same methodology, with the expected result of a moderate to strong positive correlation between level 2 state affect scores (i.e., the between-person average across episodes) and Trait Happiness scores (obtained from participants on day 1).

Long-term benefits of participating in the DRM. This was the only analysis that used data from all three participant groups, i.e., DRM, Close Control and Far Control. As a preliminary step, ANOVA and Chi-Square tests were applied to demographics to ensure there were no baseline differences across the three groups. Any significant difference between groups with more than a negligible effect size would be entered as a covariate in main analyses.

For the main analysis, we tested whether the two-day DRM protocol per se, compared with the two control groups, affected either of two self-report dependent measures: trait mindfulness (trait-FFMQ total score), and trait happiness (SHS). Note that these two trait measures were chosen to directly reflect the content of the state measures repeatedly asked in the DRM. For each trait measure, a mixed ANOVA assessed whether the mean trait score differed as a function of time (Pre-Intervention, Post-Intervention, Follow-up; measured within-person), condition (DRM, Close Control, Far Control; measured between-person), or their interaction. We expected to observe small but significant benefits for both trait measures for participants in the DRM condition (i.e. an increase from Pre-Intervention to Post-Intervention that is sustained at Follow-up; see Procedure), yet no benefits for participants in either of the two control conditions.

Results

Descriptive. We assumed the 370 participants in the total sample would be evenly distributed across conditions since attrition and exclusion rates from data cleaning were comparable in the three conditions. Therefore, the resulting discrepancy in final sample sizes between groups, particularly for the Far control (DRM n = 113; Close Control n = 107; Far Control n = 150) was unexpected. Although the quasi-randomization procedure did not unfold as planned, the sample sizes for each group was still large enough to conduct all planned analyses with integrity; hence, we proceeded without modification. Demographic information can be found in Table 2.1. Overall most participants were female (80.81%) and Asian (45.67%) with a mean age of 20.87 years (range 18-46).

In the DRM condition, 1751 total episodes were recorded across 113 participants (M = 15.50, SD = 6.42, range = 3 - 38). The median duration to complete one DRM day during the intervention protocol was 28.88 minutes. To ensure participants sufficiently remember the episode they were reporting on, the reinstantiation task asked participants to report on how much of each episode they remembered. In the DRM condition, participants rated remembering some (n = 586 episodes; 33.47%), most (n = 770 episodes; 44.00%), or all (n = 395 episodes; 22.56%)

of the episode, with no reports (0 episodes; 0%) of participants remembering none or very little of an episode. For Activity Type, most of the DRM episodes were categorized as Restful (507; 28.95%), Cognitive (439; 25.07%), or Social (350; 19.99%), followed by Household (196; 11.19%), Other (130; 7.42%), and Physical (129; 7.37%).

Table 2.1: Demographic Information for the Three Conditions

	DRM (N=113)	**Close Control (N=107)**	**Far Control (N=150)**	**Total (N=370)**
Age				
Mean (SD)	20.9 (3.02)	21.4 (3.96)	20.5 (2.82)	20.9 (3.26)
Median [Min, Max]	20.0 [18.0, 38.0]	21.0 [18.0, 46.0]	20.0 [18.0, 43.0]	20.0 [18.0, 46.0]
Sex at Birth				
Female	90 (79.6%)	92 (86.0%)	117 (78.0%)	299 (80.8%)
Male	23 (20.4%)	15 (14.0%)	33 (22.0%)	71 (19.2%)
Ethno-Racial Category				
Asian	50 (44.2%)	40 (37.4%)	79 (52.7%)	169 (45.7%)
Hispanic or Latino	26 (23.0%)	24 (22.4%)	23 (15.3%)	73 (19.7%)
White	22 (19.5%)	29 (27.1%)	24 (16.0%)	75 (20.3%)
Mixed	11 (9.7%)	10 (9.3%)	14 (9.3%)	35 (9.5%)
Black or African American	2 (1.8%)	2 (1.9%)	5 (3.3%)	9 (2.4%)
Middle Eastern or North African	1 (0.9%)	2 (1.9%)	4 (2.7%)	7 (1.9%)
Native Hawaiian or other Pacific Islander	1 (0.9%)	0 (0%)	1 (0.7%)	2 (0.5%)
Meditation Status				
Non-meditator	87 (77.0%)	87 (81.3%)	111 (74.0%)	285 (77.0%)
Past meditator	16 (14.2%)	9 (8.4%)	18 (12.0%)	43 (11.6%)
Current meditator	10 (8.8%)	11 (10.3%)	21 (14.0%)	42 (11.4%)

In the Close Control condition, 2225 total episodes were recorded across 107 participants (M = 21.19, SD = 8.76, range = 6 - 51). The median duration to complete one Close Control day during the intervention protocol was 11.11 minutes.

As an exploratory measure, we asked participants from all three groups on day 4 about the typicality of the past two days. Overall, while most participants (n = 298, 80.50%) reported that the past two days were fairly typical, a sizable proportion of the sample reported that something majorly upsetting (n = 41; 11.10%) or wonderful (n = 31; 8.38%) happened in the past two days. A chi-square test revealed no significant differences in these proportions between the three participant groups, X^2 (4, N = 370) = 2.78, p = .60.

Do the Different Components of State Mindfulness Uniquely Predict State Affect? Note that all analyses in this section use the DRM condition exclusively. The main empirical test was to assess the unique predictive effects of each facet of state mindfulness on state affect in a MLM. For this MLM, state affect was the dependent variable, the four state-FFMQ facets were entered simultaneously as predictor variables (fixed effects), and participant ID was entered as a random intercept effect (Table 2.2, *left panel*). Awareness, Nonjudging, and Observing uniquely and significantly predicted state affect, whereas Describing had no significant predictive value above and beyond the other facets. Furthermore, the strength of each significant predictor varied. The largest share of variance was explained by Awareness and Nonjudging, which uniquely explained 6% and 4% of the variance in state affect, respectively. Observing uniquely explained 0.3% of the variance in state affect.

To further substantiate these results, we added in several relevant fixed effect covariates. To ensure that the trait-FFMQ, Activity Valence, and Activity Type were suitable to be used as covariates, we first examined bivariate associations between each covariate with the state affect. Trait mindfulness (Pseudo-R^2 = .008), Activity Valence (Pseudo-R^2 = .393), and Activity Type (Pseudo-R^2 = .190) were all significantly and positively related with state affect scores (p's < .001). We further confirmed that none of the potential covariates interacted with the state-FFMQ total score in predicting state affect. Therefore, these three variables were all suitable as covariates.

In the robust model including Activity Valence and Activity Type as covariates[12] (Table 2.2, *right panel*), the effect sizes of all state-FFMQ facets did substantially decrease from the predictive model. While much of this decrease can be attributed to the relatively strong effect of Activity Valence (Pseudo-R^2 = .175), the effect of Activity Type (Pseudo-R^2 = .022) also played a role[13]. Within the Activity Type factor there were notable differences based on type of activity engaged in: Resting, Social, and Physical activities were associated with greater state affect than Household, Cognitive, or Other activities.

In the incremental model including just trait mindfulness as a covariate (Table 2.2, *middle panel)*, the effect sizes of the state-FFMQ facets did not change from the predictive model. Though the effect of trait mindfulness was statistically significant ($p < .001$), it uniquely accounted for a relatively small share of variance in the model (Pseudo-R^2 = .006).

Further Validation of the Newly Created State-FFMQ. These analyses also only used data from the DRM condition. Since we found evidence that several components of state mindfulness uniquely predicted state affect, this provided evidence of the predictive validity of the state-FFMQ, as was observed in the development of the state-FFMQ (Raynes & Dobkins, under review).

[12] Though not pre-registered, we also ensured that Activity Valence and Activity Type were sufficiently unique from each other. In a MLM, Activity Type did significantly predict Activity Valence (Pseudo-R^2 = .207, $p < .001$), but its effect was not so strong as to suggest completely overlapping constructs.

[13] Though not pre-registered, we also explored a model that used only Activity Valence as a covariate, and a model that used only Activity Type as a covariate. Though the effect of the state-FFMQ decreased in both exploratory models compared to the predictive validity model, the decrease of effects was much more notable in the model using only Activity Valence as a covariate.

Table 2.2: Variables Predicting State Affect During Daily Life Experiences

	Predictive Model				Incremental Model				Robust Model			
Predictors	*B*	*CI*	*p*	*Pseudo-R^2*	*B*	*CI*	*p*	*Pseudo-R^2*	*B*	*CI*	*p*	*Pseudo-R^2*
(Intercept)	0.58	0.49 – 0.68	**<.001**		-0.54	-1.10 – 0.02	.059		0.37	0.26 – 0.48	**<.001**	
Awareness	0.33	0.27 - 0.38	**<.001**	0.061	0.33	0.27 – 0.38	**<.001**	0.061	0.18	0.14 – 0.22	**<.001**	0.018
Observing	0.08	0.02 - 0.14	**.009**	0.003	0.08	0.02 – 0.14	**.009**	0.003	-0.01	-0.06 – 0.03	.590	0.0001
Describing	0.02	-0.05 - 0.10	.525	0.0002	0.02	-0.05 – 0.10	.525	0.0002	0.05	-0.01 – 0.11	.081	0.001
Nonjudging	0.36	0.29 - 0.43	**<.001**	0.040	0.36	0.29 – 0.43	**<.001**	0.040	0.27	0.21 – 0.32	**<.001**	0.022
Trait Mindfulness					0.02	0.01 - 0.04	**<.001**	0.006				
Activity Valence									0.35	0.32 – 0.38	**<.001**	0.175
Activity Type												0.022
Household									0.07	-0.05 – 0.19	.228	
Physical									0.36	0.22 – 0.50	**<.001**	
Other									-0.04	-0.17 – 0.10	.597	
Restful									0.36	0.26 – 0.46	**<.001**	
Social									0.38	0.27 – 0.49	**<.001**	
σ^2	0.78				0.78				0.44			
τ_{00}	0.21 $_{\text{Participant}}$				0.18 $_{\text{Participant}}$				0.22 $_{\text{Participant}}$			
ICC	0.21				0.18				0.34			
N	113 $_{\text{Participant}}$				113 $_{\text{Participant}}$				113 $_{\text{Participant}}$			
Observations	1751				1751				1751			
Marginal R^2 / Conditional R^2	0.168 / 0.344				0.195 / 0.344				0.441 / 0.631			

Note. Left Panel: Model asking if the state-FFMQ facets predict Happiness. *Middle Panel*: Model asking if adding trait mindfulness lowers the effects of the state-FFMQ. *Right Panel*: Model asking if adding Activity Valence and Activity lowers the effects of the state-FFMQ. B represents unstandardized beta coefficients. CI represent confidence intervals of the beta coefficients. The pseudo-R^2 statistic is the difference in R^2 of the full model versus a model with the relevant variable removed. For continuous variables, positive effect sizes represent that an increase in the predictor variable is associated with an increase in state affect (and vice versa for negative effect sizes). For Activity, a positive effect size means that Activity led to higher happiness than did the referent level of Cognitive Activity (and vice versa for negative effect sizes). Bolded values are statistically significant.

Construct validity of the state-FFMQ was examined in three ways. First, we tested the multidimensionality of the state-FFMQ. We compared the model fit between the predictive model above, versus an alternative model of one total state-FFMQ score. As expected, the model incorporating all four facets demonstrated superior fit statistics to the model with one total score, as evidenced by lower BIC (4755.0 versus 4818.1) and AIC (4716.8 vs. 4796.2) values. Furthermore, a chi-square difference test revealed a significant improvement in model fit when including the individual facets, $\Delta\chi^2(3) = 85.48$, $p < .001$.

Second, we tested whether the state-FFMQ was behaving in a state- rather than trait-like manner. In the incremental model above (Table 2.2, *middle panel*), the relationship between the state-FFMQ and state affect remained strong even after accounting for the effects of the trait-FFMQ. We also calculated ICC statistics to determine the amount of within-person variance of each facet during the Intervention protocol (see Table 2.3). As predicted, about half or more of the variation in all state-FFMQ facets came from within rather than between person variation. In addition, there was a wide range of within-person variance attributable to each facet (49.39% - 73.62%).

Third, we tested the discriminant sensitivity of the state-FFMQ in detecting mindful states by assessing whether the level 2 state-FFMQ total score (i.e., averaged across all episodes within a person) differs by Meditation Status, with the expectation that current and past meditators experienced more daily mindfulness on average than non-meditators. The ANOVA test revealed a significant main effect of Meditation Status on the level 2 state-FFMQ total score with a medium effect size ($F(2, 110) = 3.91$, $p = .02$, $\eta^2 = .07$). Tukey's Honestly Significant Difference post-hoc test was conducted to explore pairwise differences between Meditation Status categories. The results of the post-hoc tests revealed that non-meditators had marginally lower average state mindfulness scores than current meditators (mean difference = -0.37, p = .08) and past meditators (mean difference = -0.29, p = .11). While these results were not statistically significant, the direction of effects was as expected. No difference was observed between current and past meditators (mean difference = -0.08, p = .92).

Table 2.3: Intraclass Correlation Coefficients and Grand Mean Estimates

Variable	Between-Person Variance (%)	Within-Person Variance (%)	Grand Mean Estimate[a]
State Affect	16.45	83.55	0.58
State FFMQ Total	50.59	49.41	3.57
Awareness	26.38	73.62	3.33
Observing	44.33	55.67	3.22
Describing	48.02	51.98	3.65
Nonjudging	50.61	49.39	4.08
Activity Valence	9.58	90.42	3.74

[a] The grand mean estimate refers to the overall mean score across all episodes and participants. State Affect was scored on a sliding scale ranging from -2 to +2. State-FFMQ was scored on a sliding scale ranging from 1-5. Activity Valence was scored on a Likert scale from 1-7.

Convergent validity assessed whether each state-FFMQ facet was accurately measuring the trait-FFMQ facet from which it derived by using bivariate correlations to assess the level 2 state-FFMQ facets versus their corresponding trait-FFMQ facet score (obtained in day 1). As predicted, the aligned[14] facets of the state and trait FFMQ had medium to large effect sizes with the following values: Total score (r(111) = .51, p < .001), Awareness (r(111) = .27, p = .003),

[14] Though not pre-registered, we realized that we ought to also look at the correlations of misaligned facets, as was done in the development article of the state-FFMQ, and found two unexpected findings. First, state Awareness was marginally more strongly correlated with trait Nonjudging (r = .33) than its aligned facet (r = .27). Second, state Describing was marginally more strongly correlated with trait Observing (r = .37) and trait Nonjudging (r = .38) than with its aligned facet (r = .35). This non-replication of the convergent validity results from the development article might be due to differences in the reference point (i.e. the previous day's events, versus an immediately preceding meditation), and is an interesting area for future exploration.

Observing (*r*(111) = .41, $p < .001$), Nonjudging (*r*(111) = .52, $p < .001$), and Describing (*r*(111) = .35, $p < .001$). Also as expected, exploratory analyses revealed that there was a moderate to large relationship between level 2 state affect versus Pre-Intervention trait Happiness (*r*(111) = .44, $p < .001$).

Long-term benefits of participating in the DRM. Note that this analysis used the Total sample. As a preliminary step, ANOVA and Chi-Square tests were applied to demographics and Meditation Status (see Table 2.1 above) to check for baseline differences across the three conditions. The conditions were comparable at baseline, with no significant differences across the groups. Therefore, no covariates were added to the subsequent models. For trait mindfulness, the results of the two-way mixed ANOVA revealed a significant main effect for condition ($F(2, 1101) = 14.51$, $p < .001$). However, both the main effect for time (F(2, 1101) = 2.46, $p = .11$) and the interaction between time and condition ($F(4, 1026) = 0.80$, $p = .52$) were not significant. For trait happiness, the results of the two-way mixed ANOVA revealed a significant main effect for condition ($F(2, 1101) = 10.62$, $p < .001$). However, both the main effect for time ($F(2, 1101) = 0.44$, $p = .64$) and the interaction between time and condition ($F(4, 1101) = 0.196$, p = .94) was not close to significant[15].

Discussion

The results of the current study show that state mindfulness – captured in participants' retrospective reflections of "daily life" episodes from the previous day – predicts state affect, in particular, happiness. These findings obtained using the DRM complement those from previous studies using a different methodology, ESM (e.g., Blanke et al., 2018; Brown & Ryan, 2003;

[15] Though not pre-registered, exploratory analyses also tested whether there were any significant differences when omitting the Close Control group, and/or omitting the follow-up timepoint. Even with this more powerful test to detect any expected differences, we still did not find significant effects of time, or the interaction of time and condition, for either trait measure.

Enkema et al., 2020; Raugh et al., 2023; Snippe et al., 2015). One of the novel aspects of the current study was the use of a recently developed four facet state mindfulness questionnaire (i.e., the state-FFMQ), allowing us to investigate *which* aspects of state mindfulness are most closely tied to state affect. We found the strongest unique effects of Awareness and Nonjudging on state affect, with lesser effects for Observing and no effect of Describing, and these effects remained robust when several covariates were included. This pattern of results is strikingly similar to those obtained in the development article for the state-FFMQ (Raynes & Dobkins, under review), despite that previous study's use of a different methodology for measuring state mindfulness (i.e. retrospective reflections of an immediately preceding meditation) and a different dependent variable (state stress and anxiety). Although the observed small unique effects of state-Observing and state-Describing on state affect may appear surprising, these findings mirror previous studies showing that trait-Observing and trait-Describing (from the trait-FFMQ) show a weak relationship with affective symptoms (e.g., Carpenter et al., 2019; Mattes, 2019) and severity scores on a psychological inventory scale (Baer et al., 2006). As a possible explanation of these weak results in the context of the trait-FFMQ, it has been argued that trait-Observing (e.g., Christopher et al., 2012) and trait-Describing (e.g., Tran et al., 2013) are facets of mindfulness that may be more relevant for individuals with sufficient meditation experience. Because most of our sample had little meditation experience, the weak to insignificant results of state Observing and state Describing in the current study may therefore be attributable to characteristics of our sample and should be further explored in more diverse populations and research contexts.

Although our findings demonstrate that different facets of state mindfulness uniquely predict state affect in daily life, we must caution that the current findings cannot speak to the question of causality given their correlational nature. While our findings are consistent with the possibility that moments of heightened Awareness, Nonjudging, and Observing lead to moments of heightened happiness, the converse may instead (or also) be true, i.e., moments of heightened happiness could lead to moments of heightened Awareness, Nonjudging, and Observing (see Du et al., 2019 for an ESM study demonstrating a reciprocal relationship between state mindfulness and positive emotions). Though not the purpose of the current study, to establish a causal pathway, future experimental studies must show either that 1) inducing a change in one variable creates a change within another or 2) that a state of one variable at time 0 predicts the state of another at time 1 (noting that the time lag needs to be fairly immediate; see Mason et al., 2013). Regardless of the direction of causality, future research should continue to explore the relative importance of each component of mindfulness when measured in naturalistic settings using state, rather than trait, measures.

Further Validation of the State-FFMQ. A secondary aim of the current study was to further validate the state-FFMQ. Whereas the original validation of the state-FFMQ was in the context of formal mindfulness, the current study employed the state-FFMQ in the context of daily mindfulness. As such, the current study provides an opportunity to further validate the new scale when applied to a very different setting. Below, we discuss evidence for predictive, convergent and construct validity of the state-FFMQ.

The first type of validation criteria that was met in the current study is predictive validity, demonstrated simply by the fact our main empirical analysis found that state mindfulness predicts state affect. Second, convergent validity of the state-FFMQ is supported by the significant bivariate correlations between the level 2 state facets and their aligned trait-FFMQ facets. The predictive and convergent validity observed in the current study corroborates the findings from the development article (Raynes & Dobkins, under review).

Construct validity of the state-FFMQ is demonstrated in three ways. First, it behaves in a multidimensional fashion. In the development article (Raynes & Dobkins, under review), multidimensionality was demonstrated by comparing factor structures (i.e. a four-factor versus a single-factor model), with the four-factor model showing superior fit statistics. In the current study, we corroborate this result by showing multidimensionality through the predictive model wherein a model fit comparisons suggests that incorporating the individual facets of the state-FFMQ provides valuable information beyond what is captured by the total score alone in predicting state affect, despite the increased complexity associated with including multiple facets within a model.

Second, the state-FFMQ behaves more like a state than a trait measure. In the predictive validity model (mentioned above), adding *trait* mindfulness as a covariate had no impact on the relationship between state mindfulness and state affect, and trait mindfulness itself had a very small but significant main effect. In addition to corroborating an analogous validation result in the development article (Raynes & Dobkins, under review), this finding is fascinating because it means that compared to generalized baseline levels of trait mindfulness, daily fluctuations of state mindfulness facets bear greater relevance in predicting daily experiences of state affect. The demonstration of the state-like nature of the state-FFMQ is further bolstered by the results of the ICC analysis, which revealed substantial moment-to-moment variation in all state-FFMQ facets. Future research might consider taking a more granular approach to measuring within-person variation by assessing the variance of individual items within versus between people, which is a concept similar to Generalizability theory (see Medvedev et al., 2017; Truong et al., 2020).

Third, the state-FFMQ demonstrates discriminant sensitivity in detecting mindful states. Even over the course of just two days, the state-FFMQ detected greater average (level 2) daily mindfulness levels for current and past, than non, meditators. Larger effects may be expected amongst samples with more variance in the Meditation Status groupings, and if the length of administration were longer than two days. Note that this form of construct validity can be applied either to the type of participant (i.e. by meditation status, as was done in the current study and the development article of the state-FFMQ), or by an experimental condition (e.g., in reference to a meditation versus a control condition, as was done in the development article of the state-FFMQ). In both the current study and the development article of the state-FFMQ (Raynes & Dobkins, under review), the results of this analysis by type of participant provide evidence for discriminant sensitivity.

Are there Long-Term Benefits of Participating in the DRM? A third goal of the current study was to ask whether there are long-term benefits of participating in the DRM. Here, we found that none of the three participant groups demonstrated improvements in trait mindfulness or trait happiness. This null result may simply reflect that two days is not enough time to significantly alter trait mindfulness or happiness. In fact, some previous studies suggest that improving trait mindfulness or trait happiness takes intensive time and effort and may not even produce lasting effects, even in targeted intervention studies (e.g., see Seligman et al., 2005; Visted et al., 2015). Still, given that other daily reflecting interventions, such as journaling about one's days, can demonstrate long-term psychological benefits (e.g., Dimitroff et al., 2016; Keech & Coberly-Holt, 2021; Smyth et al., 2018), we believe the current DRM design, which asks people to reflect on their days within the context of noticing mindful moments, could produce long-term benefits if the protocol duration were longer.

Limitations and Considerations. On a final note, we discuss three general considerations that apply to future work employing the DRM to assess daily mindfulness. First, it is important to address the trustworthiness of DRM data in general, given that the method involves asking participants to make retrospective reports on what they remember internally experiencing during events of variable duration from the previous day. The fact that the current, and previous, DRM studies find significant associations between their variables of interest suggest that the DRM method must be reliable to some degree. Whereas most previous DRM studies have found that the method accurately captures recalled *affective* states by finding consistent results when assessed across modalities (e.g., Diener & Tay, 2014; Kahneman et al., 2004; Krueger et al., 2004; Ludwigs et al., 2019; Schwarz et al., 2009), what is novel about the current study is that it provides preliminary evidence that the DRM reliably captures recalled *mindful* states. Still, it is possible that individuals may not be capable of accurately discerning every facet of state mindfulness across every previous-day episode. The boundary conditions outlining when it is valid to make such self-evaluations about state mindfulness is an unexplored research question. One way to investigate this would be to include a Don't know or Not Applicable (DK/NA) response option to the state mindfulness questionnaire. This was not utilized in the current study due to the risk that participants may overuse this option as a means of speeding through the survey with insufficient effort. However, future research can explore this approach to see whether the relative frequency of DK/NA responses are unequal across state mindfulness items, activity types and activity valences, or even types of people (e.g., based on baseline trait mindfulness scores or meditation status). This approach may be particularly useful in DRM and EOD methods since previous work has shown that the valence of an event can affect the accuracy of general memory recall (see Colombo et al., 2024), and this likely extends to specific memory recalls of mindful states.

Note that further research will be needed to elucidate the effects of various operationalizations of “valence” on state mindfulness in different contexts. Specifically, we recommend accounting for activity valence (e.g., Colombo et al., 2024; see also "daily stressors" as measured in Miller et al., 2024), thought valence (e.g., Banks et al., 2016; Gross et al., under review; Mills et al., 2021; Welz et al., 2018), or other aspects of thought *content* related to valence such as time orientation or relevance of thoughts (e.g., Poerio et al., 2013), depending on the research design.

A second question is whether the DRM can be employed for multiple consecutive days without overburdening participants. Although a handful of studies have conducted consecutive multi-day DRM designs (e.g., Dockray et al., 2010; Ludwigs et al., 2019), the results of the current study show that having participants complete the DRM for two consecutive days is feasible and can be further explored in subsequent research. We found that participants reported an average of about 15 episodes per day, which is similar to the number of average episodes per day reported for previous DRM studies (e.g., Kahneman et al., 2004; Ludwigs et al., 2019; Schneider et al., 2020). This suggests that participants were actively engaged with both days of the DRM. We also found that it took participants a median of under 30 minutes to complete each DRM day. This is substantially less than the 45-75 minute duration estimate provided by Kahneman et al. (2004), which has been cited as a reason for selecting a small subset of randomly selected episodes for participants to respond to DRM questions about (e.g., Hudson et al., 2020) rather than all episodes as in the current study. As the DRM took much shorter than expected, it may therefore feasible to ask participants about all episodes reinstantiated, and therefore get a more representative dataset, without overburdening participants.

A third question is whether researchers should take into account the typicality of recalled episodes. Like many ILD studies, the current research intended on capturing events that frequently occur for most participants in daily life. To test this, we simply asked participants whether the previous two days were typical for them or not, and found that about 20% did *not* have a “typical” experience, roughly half of which were reported to be upsetting versus wonderful. Though we had no a priori estimate, this was higher than expected and could have impacted our results. It would be ideal for future work to capture “typicality”, possibly even at the level of individual episodes, so that researchers can explore the role of this construct in greater detail.

Conclusion. The current study demonstrates that different components of state mindfulness uniquely contribute to state affect in daily life. At an applied level, these finding suggest that integrating mindfulness into daily life, particularly through Awareness and Nonjudgment, might provide a pragmatic approach to enhancing wellbeing, one that is perhaps more accessible to people than formal mindfulness practices like meditation (see Grabovac et al., 2011). Moreover, the positive relationship we see between daily mindfulness and happiness may have implications for *other* psychological and physiological constructs. As such, future studies might explore the predictive effects of the state-FFMQ on states of arousal, cognitive performance, or physical symptoms, to name a few.

In addition to these empirical findings, the current study provides further validation for the state-FFMQ. Whereas the original validation of the state-FFMQ was in the context of formal mindfulness, the current study employed the state-FFMQ in the context of daily mindfulness. Like the development article, the current study demonstrates predictive, construct, and convergent validity. As such, the state-FFMQ is currently the only multidimensional state mindfulness measure that has undergone, and passed, several validity tests when used in formal and daily mindfulness contexts. While further studies are needed to replicate these results - especially amongst more diverse populations given our sample was overrepresented by female and Asian participants with limited meditation experience - the initial findings reported here for the state-FFMQ are promising.

Acknowledgements

Chapter 2 has been submitted for publication of the material as it may appear in Mindfulness, 2024, Raynes, Stephen; Dobkins, Karen, Springer. The dissertation author was the primary researcher and author of this paper.

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

Bamberger, K. T. (2016). The Application of Intensive Longitudinal Methods to Investigate Change: Stimulating the Field of Applied Family Research. *Clinical Child and Family Psychology Review*, *19*(1), 21–38. https://doi.org/10.1007/s10567-015-0194-6

Banks, J. B., Welhaf, M. S., Hood, A. V. B., Boals, A., & Tartar, J. L. (2016). Examining the role of emotional valence of mind wandering: All mind wandering is not equal. *Consciousness and Cognition*, *43*, 167–176. https://doi.org/10.1016/j.concog.2016.06.003

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

Blanke, E. S., & Brose, A. (2022). Multidimensional State Mindfulness Questionnaire (MSMQ). In O. N. Medvedev, C. U. Krägeloh, R. J. Siegert, & N. N. Singh (Eds.), *Handbook of Assessment in Mindfulness Research* (pp. 1–13). Springer International Publishing. https://doi.org/10.1007/978-3-030-77644-2_24-1

Blanke, E. S., Riediger, M., & Brose, A. (2018). Pathways to happiness are multidirectional: Associations between state mindfulness and everyday affective experience. *Emotion*, *18*(2), 202–211. https://doi.org/10.1037/emo0000323

Bolger, N., & Laurenceau, J.-P. (2013). *Intensive Longitudinal Methods: An Introduction to Diary and Experience Sampling Research*. Guilford Press.

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

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

Bylsma, L., Taylor-Clift, A., & Rottenberg, J. (2011). Emotional Reactivity to Daily Events in Major and Minor Depression. *Journal of Abnormal Psychology*, *120*, 155–167. https://doi.org/10.1037/a0021662

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., & Baitmangalkar, 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

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

Colombo, D., Wrzesien, M., & Baños, R. M. (2024). Recalling Tough Times: Exploring the Effect of Acceptance on Autobiographical Memory in Daily Life. *Mindfulness*. https://doi.org/10.1007/s12671-024-02301-0

Diener, E., & Tay, L. (2014). Review of the Day Reconstruction Method (DRM). *Social Indicators Research*, *116*(1), 255–267. https://doi.org/10.1007/s11205-013-0279-x

Dimitroff, L., Sliwoski, L., O'Brien, S., & Nichols, L. (2016). Change your life through journaling—The benefits of journaling for registered nurses. *Journal of Nursing Education and Practice*, *7*(2), 90–98. https://doi.org/10.5430/jnep.v7n2p90

Dockray, S., Grant, N., Stone, A. A., Kahneman, D., Wardle, J., & Steptoe, A. (2010). A Comparison of Affect Ratings Obtained with Ecological Momentary Assessment and the Day Reconstruction Method. *Social Indicators Research*, *99*(2), 269–283. https://doi.org/10.1007/s11205-010-9578-7

Du, J., An, Y., Ding, X., Zhang, Q., & Xu, W. (2019). State mindfulness and positive emotions in daily life: An upward spiral process. *Personality and Individual Differences*, *141*, 57–61. https://doi.org/10.1016/j.paid.2018.11.037

Enders, C. K., & Tofighi, D. (2007). Centering predictor variables in cross-sectional multilevel models: A new look at an old issue. *Psychological Methods*, *12*(2), 121–138. https://doi.org/10.1037/1082-989X.12.2.121

Enkema, M. C., McClain, L., Bird, E. R., Halvorson, M. A., & Larimer, M. E. (2020). Associations Between Mindfulness and Mental Health Outcomes: A Systematic Review of Ecological Momentary Assessment Research. *Mindfulness*, *11*(11), 2455–2469. https://doi.org/10.1007/s12671-020-01442-2

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

Grabovac, A. D., Lau, M. A., & Willett, B. R. (2011). Mechanisms of Mindfulness: A Buddhist Psychological Model. *Mindfulness*, *2*(3), 154–166. https://doi.org/10.1007/s12671-011-0054-5

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

Henwood, A., Guerreiro, J., Matic, A., & Dolan, P. (2022). The duration of daily activities has no impact on measures of overall wellbeing. *Scientific Reports*, *12*(1), 514. https://doi.org/10.1038/s41598-021-04606-9

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

Hudson, N. W., Anusic, I., Lucas, R. E., & Donnellan, M. B. (2020). Comparing the Reliability and Validity of Global Self-Report Measures of Subjective Well-Being With Experiential Day Reconstruction Measures. *Assessment*, *27*(1), 102–116. https://doi.org/10.1177/1073191117744660

Hudson, N. W., Lucas, R. E., & Donnellan, M. B. (2017). Day-To-Day Affect is Surprisingly Stable: A 2-Year Longitudinal Study of Well-Being. *Social Psychological and Personality Science*, *8*(1), 45–54. https://doi.org/10.1177/1948550616662129

Kabat-Zinn, J. (1982). An Outpatient Program in Behavioral Medicine for Chornic 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

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

Keech, K., & Coberly-Holt, P. (2021). *Journaling for Mental Health* (pp. 39–44). https://doi.org/10.4018/978-1-7998-4477-8.ch005

Killingsworth, M. A., & Gilbert, D. T. (2010). A Wandering Mind Is an Unhappy Mind. *Science*, *330*(6006), 932–932. https://doi.org/10.1126/science.1192439

Kim, J., Kikuchi, H., & Yamamoto, Y. (2013). Systematic comparison between ecological momentary assessment and day reconstruction method for fatigue and mood states in healthy adults. *British Journal of Health Psychology*, *18*(1), 155–167. https://doi.org/10.1111/bjhp.12000

Knabe, A., Rätzel, S., Schöb, R., & Weimann, J. (2010). Dissatisfied with Life but Having a Good Day: Time-use and Well-being of the Unemployed. *The Economic Journal*, *120*(547), 867–889. https://doi.org/10.1111/j.1468-0297.2009.02347.x

Krueger, A., Schkade, D., Schwarz, N., & Stone, A. (2004). *The Day Reconstruction Method (drm): Instrument Documentation*.

Lucas, R. E., Wallsworth, C., Anusic, I., & Donnellan, M. B. (2021). A direct comparison of the day reconstruction method (DRM) and the experience sampling method (ESM). *Journal of Personality and Social Psychology*, *120*(3), 816–835. https://doi.org/10.1037/pspp0000289

Ludwigs, K., Henning, L., & Arends, L. R. (2019). Using the Day Reconstruction Method—Same Results when Used at the End of the Day or on the Next Day? *International Journal of Community Well-Being*, *2*(1), 61–73. https://doi.org/10.1007/s42413-019-00017-x

Lyubomirsky, S., & Lepper, H. S. (1999). A Measure of Subjective Happiness: Preliminary Reliability and Construct Validation. *Social Indicators Research*, *46*(2), 137–155. https://doi.org/10.1023/A:1006824100041

Mason, M. F., Brown, K., Mar, R. A., & Smallwood, J. (2013). Driver of discontent or escape vehicle: The affective consequences of mindwandering. *Frontiers in Psychology*, *4*. https://doi.org/10.3389/fpsyg.2013.00477

Mathieu, J. E., Aguinis, H., Culpepper, S. A., & Chen, G. (2012). Understanding and estimating the power to detect cross-level interaction effects in multilevel modeling. *Journal of Applied Psychology*, *97*(5), 951–966. https://doi.org/10.1037/a0028380

Mattes, J. (2019). Systematic Review and Meta-Analysis of Correlates of FFMQ Mindfulness Facets. *Frontiers in Psychology*, *10*. https://www.frontiersin.org/articles/10.3389/fpsyg.2019.02684

Medvedev, O., Krägeloh, C., Narayanan, A., & Siegert, R. (2017). Measuring Mindfulness: Applying Generalizability Theory to Distinguish between State and Trait. *Mindfulness*, *8*. https://doi.org/10.1007/s12671-017-0679-0

Miller, R. L., Shomaker, L. B., Prince, M. A., Haddock, S., Rzonca, A., Krause, J. T., Zimmerman, T., Lavender, J. M., Sibinga, E., & Lucas-Thompson, R. G. (2024). Momentary effects of life stressors on mindfulness and emotion regulation difficulties among adolescents exposed to chronic stressors. *Stress and Health*, *n/a*(n/a), e3414. https://doi.org/10.1002/smi.3414

Mills, C., Porter, A. R., Andrews-Hanna, J. R., Christoff, K., & Colby, A. (2021). How task-unrelated and freely moving thought relate to affect: Evidence for dissociable patterns in everyday life. *Emotion*, *21*(5), 1029–1040. https://doi.org/10.1037/emo0000849

Nezlek, J. B. (2012). Multilevel modeling for psychologists. In H. Cooper, P. M. Camic, D. L. Long, A. T. Panter, D. Rindskopf, & K. J. Sher (Eds.), *APA handbook of research methods in psychology, Vol 3: Data analysis and research publication.* (pp. 219–241). American Psychological Association. https://doi.org/10.1037/13621-011

Oerlemans, W. G. M., & Bakker, A. B. (2013). Capturing the Moment in the Workplace: Two Methods to Study Momentary Subjective Well-Being. In A. B. Bakker (Ed.), *Advances in Positive Organizational Psychology* (Vol. 1, pp. 329–346). Emerald Group Publishing Limited. https://doi.org/10.1108/S2046-410X(2013)0000001017

Oerlemans, W. G. M., Bakker, A. B., & Veenhoven, R. (2011). Finding the Key to Happy Aging: A Day Reconstruction Study of Happiness. *The Journals of Gerontology Series B: Psychological Sciences and Social Sciences*, *66B*(6), 665–674. https://doi.org/10.1093/geronb/gbr040

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

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

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

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

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

Schneider, S., Junghaenel, D. U., Gutsche, T., Mak, H. W., & Stone, A. A. (2020). Comparability of Emotion Dynamics Derived From Ecological Momentary Assessments, Daily Diaries, and the Day Reconstruction Method: Observational Study. *Journal of Medical Internet Research*, *22*(9), e19201. https://doi.org/10.2196/19201

Schneider, S., & Stone, A. A. (2016). Ambulatory and diary methods can facilitate the measurement of patient-reported outcomes. *Quality of Life Research*, *25*(3), 497–506. https://doi.org/10.1007/s11136-015-1054-z

Schwarz, N., Kahneman, D., & Xu, J. (2009). Global and Episodic Reports of Hedonic Experience. In R. Belli, F. Stafford, & D. Alwin, *Calendar and Time Diary* (pp. 156–174). SAGE Publications, Inc. https://doi.org/10.4135/9781412990295.d15

Seligman, M. E. P., Steen, T. A., Park, N., & Peterson, C. (2005). Positive Psychology Progress: Empirical Validation of Interventions. *American Psychologist*, *60*(5), 410–421. https://doi.org/10.1037/0003-066X.60.5.410

Smyth, J. M., Johnson, J. A., Auer, B. J., Lehman, E., Talamo, G., & Sciamanna, C. N. (2018). Online Positive Affect Journaling in the Improvement of Mental Distress and Well-Being in General Medical Patients With Elevated Anxiety Symptoms: A Preliminary Randomized Controlled Trial. *JMIR Mental Health*, *5*(4), e11290. https://doi.org/10.2196/11290

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*. https://doi.org/10.1007/s12671-014-0354-7

Stone, A. A., Schwartz, J. E., Schkade, D., Schwarz, N., Krueger, A., & Kahneman, D. (2006). A population approach to the study of emotion: Diurnal rhythms of a working day examined with the day reconstruction method. *Emotion*, *6*(1), 139–149. https://doi.org/10.1037/1528-3542.6.1.139

Stone, A. A., & Shiffman, S. (1994). Ecological momentary assessment (EMA) in behavorial medicine. *Annals of Behavioral Medicine*, *16*(3), 199–202. https://doi.org/10.1093/abm/16.3.199

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

Visted, E., Vøllestad, J., Nielsen, M. B., & Nielsen, G. H. (2015). The Impact of Group-Based Mindfulness Training on Self-Reported Mindfulness: A Systematic Review and Meta-analysis. *Mindfulness*, *6*(3), 501–522. https://doi.org/10.1007/s12671-014-0283-5

Welz, A., Reinhard, I., Alpers, G. W., & Kuehner, C. (2018). Happy Thoughts: Mind Wandering Affects Mood in Daily Life. *Mindfulness*, *9*(1), 332–343. https://doi.org/10.1007/s12671-017-0778-y

Chapter 3 When is a wandering mind unhappy? The role of thought valence

Abstract

The current research represents one of the first attempts to investigate how various thought qualities that naturally fluctuate across attention states (i.e., mind wandering vs present-focused attention) impact mood. Of specific interest was whether thought valence may mediate previously reported effects of attention state on mood. To examine this, an experience sampling methodology (ESM) was used to capture participants (N=337) attention state (present or mind wandering), thought valence, and mood six times per day for seven days during daily life. Participants further indicated the form of their thoughts (e.g., inner speech), as well as their clarity and interestingness. This design allowed for a conceptual replication and expansion of Killingsworth and Gilbert (2010) in which it was observed that mind wandering leads to relatively poorer mood compared to present-focused attentional states, with the poorest mood for negatively valenced wandering thoughts. Unlike their study, however, we inquired about thought valence for both mind wandering and present moments, while also restricting analyses to prompts that contained inner speech to achieve a clearer comparison. Our findings revealed that the relationship between attention state and mood is substantially accounted for by thought valence, while interestingness and clarity further provided significant, albeit much weaker, indirect effects on mood. Exploratory analyses suggested that participant age may moderate the relationship between attention state and mood. Overall, these findings suggest that the commonly reported detrimental impact of mind wandering on mood may largely be accounted for by certain confounding variables.

Introduction

Throughout the day, thoughts naturally ebb and flow from the present moment. Even when attempting to engage in a given task (such as reading this paper), unrelated thoughts will spontaneously arise and interrupt our focus (Mooneyham & Schooler, 2013). Commonly referred to as *mind wandering*, such thoughts are characterized as being internally-generated and decoupled from the current situation (Antrobus, 1968; Giambra, 1989). The types of thoughts one has while mind wandering can vary widely – from pleasant musings to uncomfortable or intrusive ruminations. Despite such variety in thought content, a growing consensus suggests that, on average, mind wandering is accompanied by relatively more negative mood states when compared to present-focused states (e.g., Killingsworth & Gilbert, 2010; Franklin et al., 2013; Ruby et al., 2013; Wilson et al., 2014; see Kam et al., 2024 for meta-analysis).

This notion first gained widespread traction following a seminal study by Killingsworth and Gilbert (2010) in which the authors concluded that *a wandering mind is an unhappy mind.* Using an Experience Sampling Methodology (ESM), the influence of attention state (i.e., mind wandering vs. not mind wandering") on mood was investigated in a large and diverse sample. In their study, participants responded to prompts on their smartphone 1-3 times a day as they went about their daily lives. For each prompt, they reported their *mood* ("How are you feeling right now?", from a scale of 0 = very bad to 100 = very good) and their *attention state* ("Were you thinking about something other than what you were currently doing?") with four possible response options; "yes", about something pleasant; "yes", about something neutral; "yes", about something negative, or "no". The results revealed that participants were overall less happy when they were mind wandering (vs. not mind wandering, which we hitherto refer to as being "present"). In addition (and perhaps unsurprisingly), Killingsworth and Gilbert also reported that when participants were mind wandering their *thought valence* affected their mood, with the lowest happiness scores occurring when thoughts were about something negative (vs. something neutral or positive). In fact, when participants were mind wandering about something positive, their happiness was on par with being present.

Critically, although Killingsworth and Gilbert asked participants to report thought valence when they were mind wandering, they did not ask about participants' thought valence when their attention was present[16]. Because of this omission, their findings leave open an intriguing possibility; perhaps people are happier when they are present (vs. mind wandering) *because* their thought valence is more positive during present moments (vs. mind wandering). Stated differently, the effect of attention state on mood may be *accounted for* by thought valence; this is the main question of the current study.

The notion that thought valence may account for the relationship between attention state and concurrent mood is supported by initial evidence from previous studies showing relationships between *pairs* of these three variables: attention state, thought valence and mood (although, to our knowledge, no study has reported on the relationship between all three variables at once). Several studies have replicated the effect of *attention state* on *concurrent mood*. Using a naturalistic approach similar to that used in Killingsworth and Gilbert, a reliable pattern has emerged: the higher the degree of task-unrelatedness, the poorer individuals' concurrent mood (Franklin et al., 2013; Mills et al., 2021; Theimann et al., 2023; Hobbiss et al., 2019).

On the other hand, the association between *thought valence* and *concurrent mood* remains underexplored. Using an experience-sampling procedure, Spronken et al. (2016)

[16] It is a bit curious that the authors provided no rationale for not asking about thought valence for present moment experiences. One possibility is that they assumed that present moments do not contain thoughts, and in particular, do not contain inner speech, which is a particular type of inner experience. We return to the empirical benefits of specifically asking participants if they experience inner speech below.

assessed momentary mood (“How happy do you feel at the moment?’, from 0 = not happy at all to 100 = very happy) and thought valence (‘Was your thought negative, neutral, or positive?’, from -5 = very negative to +5 = very positive) finding a positive association. In contrast, a meta-analytic review by Kam et al. (2024) aggregated data across mind wandering studies finding that *negative* thoughts tend to have a larger impact on mood than *positive* thoughts. Collectively, these results indicate a significant role of thought valence and mood. Nevertheless, it is important to note that the meta-analytic review on thought valence was limited to mind wandering data (i.e., not data across attentional states). Additionally, few of the included studies used experience-sampling, and the results were combined across a wide range of mood measures, from state-specific to more stable life satisfaction measures, captured concurrently, retrospectively, or subsequently to mind wandering episodes.

Similarly, few research paradigms have examined the relationship between *attention state* and *thought valence*. Using a laboratory-based paradigm, Marchetti and colleagues (2012) reported no association between the task relatedness of thoughts (i.e., on- vs. off-task thoughts) and thought valence (measured on continuous scales). However, the lack of a difference is not surprising given that such laboratory-based studies employ intentionally mundane tasks in order to induce mind wandering, and therefore “on-task” thoughts are unlikely to be less negative than mind wandering thoughts (see Mason et al., 2013 for further discussion). This is, in fact, a general drawback to studying the effects of attention state on thought valence (or mood) in tightly controlled, laboratory-based studies; they lack the ecological validity necessary for examining the relationship between all three variables as they naturally occur in each moment.

Correlational research does however show that mind wandering tends to center around personal concerns (e.g., Klinger, 1977; Mason et al., 2009), which are presumably often negative. While this may suggest a link between negative thoughts and mind wandering, it does not provide evidence that thought valence is *more* negative during mind wandering vs. present moment experiences. Indeed, even studies that measured thought valence in real-world settings, have not reported whether valence *differs* as a function of attention state (e.g., Welz et al., 2018; Poeria et al., 2013). For this reason, whether there exists a relationship between *attention state* and *thought valence* in naturalistic settings is still an open research question.

The main goal of the current study was therefore to investigate, in a naturalistic setting, whether thought valence is more negative when people are mind wandering vs. present (path a in Figure 3.1), which allowed us to ask whether the effect of *attention state* on *concurrent mood* (path c in Figure 3.1) is accounted for by *thought valence* (path a*b in Figure 3.1). To address this, we conducted a conceptual replication of Killingsworth and Gilbert (2010) with the main modification being that participants were asked to report their thought valence for *both* mind wandering and present moments. As the focus of the present study was to explore the concurrent relationships between attention state, thought valence, and mood– rather than establishing causal links– all three variables were considered at the same time. Our analysis aimed to see how the relationship between attention state and mood shifts when valence is included in the model, therefore we examined the mediation model shown in Figure 3.1. While mediation analysis typically suggests causality, in this case, it was employed solely to quantify the proportion of the attention state's main effect that could be explained by the indirect effect of valence.

A second key modification from the Killingworth and Gilbert study is that participants were asked to report on the *nature* of their thought, specifically, whether it contained speech (which we refer to as "Inner Speech") or not (which we refer to as "Inner Experience"; the nature of which was explored with further questions, see *Methods*). For the purpose of the current study, we restricted our analyses to prompts that contained Inner Speech, as we reasoned that asking participants to report thought valence is more straightforward in moments that do (vs. do not) contain inner speech (see *Methods* for our approach). However, all of the results were quite similar regardless of whether the prompts contained Inner Speech or not (all results using No Inner Speech prompts are available by request from the authors).

A secondary goal of the current study was to explore various variables that might influence the relationship between attention state and mood, including additional thought qualities (clarity and interestingness of thoughts), and potential moderating variables (demographics and trait measures). Specifically, our preliminary data indicated associations between Current Activity, Clarity, and Interestingness with both Attention State and Mood. Additionally, previous research suggests that interest has positive affective outcomes in cases of mind wandering (Franklin et al., 2013). As the central aim of this study was to clearly determine the role of thought valence in the connection between attention state and mood, these variables were primarily included as covariates in our analysis. Nonetheless, we briefly address their potential unique contributions in the Discussion section. Furthermore, in line with a robust body of research that demonstrates age's effect on the relationship between attentional states and thought valence (Welhaf et al., 2024), we also examined age as a moderating factor.

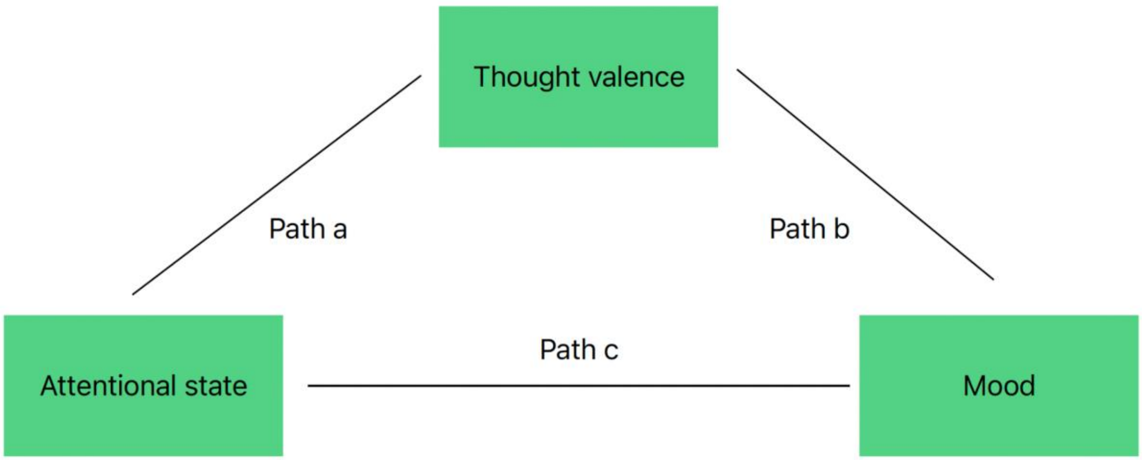


Figure 3.1: Proposed Model Describing the Association between Attention State on Mood via Thought Valence (path a*b)

Methods

Participants.

Participants were recruited for this online study through UC San Diego's (UCSD) research subject pool run by the Department of Psychology from 2022 - 2023. Participants consisted of undergraduate students who were compensated with course credit. We aimed, and were successful, at reaching a total number that was at least as large as the sample size used in our pilot studies (n = 389), as our pilot group was large enough to yield significant effects for our main hypotheses[17]. After data cleaning (described below), 337 participants, ages 18 to 44 years (*M* = 20.82, *SD* = 3.24) were retained for analysis. Most participants were female (74.78%), followed by male (24.33%) and other (0.89%). The ethnoracial makeup was predominantly Asian (48.40%), followed by Hispanic/Latino (19.00%), White or Caucasian (18.70%), Mixed (7.72%), Black or African American (1.78%), Middle Eastern or North African (1.78%), First

[17] Data from the pilot group are not presented here as the current study made some significant changes in how the prompts were delivered.

Nation, Indigenous American, Native Hawaiian or other Pacific Islander (1.18%), and Prefer not to say (1.48%). This study was approved by the IRB committee at UCSD, and all participants gave their informed consent before participating.

Procedure and Measures.

There were three parts to the study. Part 1 consisted of sending a Qualtrics link to participants, asking them to answer demographic questions (described above), validity questions (see Data Cleaning, below), a trait Mindfulness questionnaire, and a personality questionnaire. For trait mindfulness we used the 15-item Five Facet Mindfulness Questionnaires (FFMQ; Baer et al., 2006; Baer et al., 2012), which captures the following dimensions of mindfulness; Observing, Describing, Acting with Awareness, Nonjudging, and Nonreactivity. Each facet has been shown to have adequate to good internal consistency, with alpha scores ranging from .75 to .91 (Baer et al., 2006). To measure core dimensions of personality, we used the 60-item NEO Five-Factor Inventory (NEO-FFI; Costa & McCrae, 1992), which measures personality across five domains; Openness, Conscientiousness, Extraversion, Agreeableness, and Neuroticism. Each of these five dimensions has adequate internal consistency, with alpha scores ranging from .68 to .86 (Costa & McCrae, 1992). These two measures were used to determine whether the relationship between attention state and mood was moderated by these trait variables (i.e., a cross-level interaction; see Exploratory analyses in *Results*). At the end of Part 1, participants received detailed instructions on how to download and use the ExpiWell app (downloaded on a mobile device) for the ESM portion of the study.

Part 2 began the following day. First, participants received an instructional message from ExpiWell. The instructional message asked them not to respond to prompts if they were driving (for safety reasons), or if they were in class (so as to not disrupt their class time). It also introduced them to what to expect in the prompts by showing them a flowchart of the questions indicating how some questions depend on the participant's previous response.

Six ESM prompts were sent each day for the following seven days. The prompts were distributed at semi-random times from 9am to 9pm, with one prompt randomly presented within each two hour block, i.e. the first daily prompt in the morning (9am – 10:59am), the next at 11am – 12:59pm, and so on. If a participant did not initially respond to a prompt, the Expiwell app sent a notification every 15 minutes for one hour (or until the participant responded). Answering the items for each prompt (described below) took approximately 2 minutes.

Each prompt began with the following honesty prime, "Please be honest about your experience, it's really important to us. Thanks!". Participants were then asked to report the experience they were having immediately before the prompt with text that read as follows: "In the moment right before you responded to this prompt…"). Provided below, in the order they were presented to the participant, are the internal labels (*italicized*) we give to each variable (which was not seen by the participant) and the related questions they responded to.

1) *Attention State.* "In the moment right before the prompt….." 1) *PRESENT*: "My attention was related to my current activity, immediate surroundings, or inner experience" or 2) *MIND WANDERING*: My attention was NOT related to my current activity, or immediate surroundings".
2) *Mood.* "How are you feeling?", on a 7-point scale with three anchor points -3 = Very bad, 0 = Neutral , +3 = Very good.
3) *Thought Nature.* "Which best describes what you were experiencing?: 1) *INNER SPEECH*: "I was experiencing inner speech (talking to myself internally)", or 2) *INNER*

EXPERIENCE: "I was NOT experiencing inner speech (NOT talking to myself internally)".

If *inner experience* was selected, participants were further asked "Which of these were you primarily experiencing: 1) BODY: "I was experiencing body sensations, e.g. hunger", 2) EMOTION: "I was experiencing emotions, e.g. sadness", 3) ENVIRONMENT: "I was noticing the environment, e.g. looking at the trees", 4) MUSIC/SOUNDS: "I was listening to music/podcast, e.g. with my earphones in, or imagining music/sounds", 5) VISUAL IMAGERY: "I was experiencing visual imagery, e.g. imagining my dog", 6) ANOTHER PERSON: "I was experiencing another person, e.g. holding hands", 7) BLANK MIND : "I was not thinking about anything at all– my mind was completely blank", 8) OTHER: 'Please type in". *Note that data from these prompts are not analyzed in the current paper.*

The items below were asked of all participants, with the only difference being whether *inner speech* or *inner experience* was inserted in the prompt, which was dependent on the response to this *Thought Nature* question.

4) *Clarity.* "What was the clarity of your inner speech/inner experience?" on a 7-point scale with three anchor points: 1 = Not at all clear, 3 = moderately clear, 7 = Extremely clear.
5) *Thought Valence.* "What was the valence of your inner speech/inner experience? If a stranger saw the content of your inner speech/inner experience (i.e., just the actual words or the actual experience), how would they rate it?", on a 7-point scale with three anchor points: -3 =Very Negative, 0 = Neutral, +3 = Very Positive. *Note that this item was*

meant to get a more "objective" measure of the content of the inner speech, regardless of how the participant reacts to the thought (see next item).

6) *Reactivity.* "How much did you feel reactive to your inner speech/inner experience?", on a 7-point scale with three anchor points: 1 = No reaction at all, 3 = Moderate reaction, 7 = Extreme reaction. *Note that data from this item are not analyzed in the current paper.*
7) *Interestingness.* "How interesting was the inner speech/inner experience?", on a 7-point scale with three anchor points: 1 = Not at all interesting, 3 = Moderately interesting, 7 = Extremely Interesting.
8) *Current Activity.* "Whether in-person or online, what type of activity were you primarily doing? (select one)": 1) "Social activity, but NOT engaged in conversation (e.g., being around family, friends, or peers but not speaking; listening to others at a party or group outing), 2) Physical activity (e.g., exercising, sports, walking, bicycling, hiking), 3) Restful activity (e.g., eating, resting, taking a nap, doing nothing, reading for fun, watching TV or videos, browsing the internet or social media), 4) Household activity (e.g., preparing meals, grocery shopping, household finances, cleaning or other chores), 5) Cognitive activity (e.g., studying, homework, attending lecture, learning something new, puzzle solving, 6) Other activity (text entry)." *Note that these six activities are adapted from Oerlemans et al. (2011).*

Part 3 of this study involved sending a Qualtrics link to the participant after their last ESM day, in which they answered validity questions (see Data Cleaning, below).

Data Cleaning.

Participants were excluded from the study for the following reasons. *First*, if they did not complete all three parts of the study. Two hundred and nine participants were excluded for this reason. Second, participants were excluded if, in the validity question (in either Part 1 or 3), they revealed themselves to have not taken the study seriously. The validity question response options ranged from 1 = “I read all instructions and questions carefully, and answered honestly to the best of my ability” to 4 = “I tried to finish this as quickly as possible and did not read most of the questions or instructions, or I did not answer honestly”, and participants were excluded if they selected options 3 or 4. Three participants were excluded for this reason. Third, participants were excluded if they did not pass the two attention checks that were interspersed in the surveys of Part 1 and 2 (e.g., “If you are paying attention to this survey, please select [blank]”). Nine participants were excluded for this reason. Fourth, because our analyses were restricted to prompts for which participants indicated experiencing inner speech (see above), participants were excluded if their data did not include a single ESM prompt of this nature. Twelve participants were excluded for this reason. Finally, 51 participants were excluded for failing to correctly input the same participant ID number across the three parts. This resulted in a total of 337 participants left for analysis.

Data Analysis.

The main analyses in this study employed multilevel models (MLM) since the data have a natural two-level structure, where prompts collected over time are nested within individuals. All analyses were computed using R (Version 3.6.2; R Core Team, 2019) and the R-package lme4 (v1.1-27.1; Bates et al., 2015) with a maximum likelihood method of estimation. Using Type III sum of squares MLM’s, the dependent variable was *Mood*, the main predictor variable was *Attention State* (entered as a fixed effect, contrast coded as Mind Wandering = -1 and

Present = +1), with participant ID entered as a random intercept and prompt as the unit of analysis. Stated differently, the within and between subject variance of the dependent variable was partitioned by fitting random intercept terms for each participant and forcing a fixed slope. Prior to analysis, all continuous level 1 variables were person-mean centered, sometimes referred to as "centering-within-clustering", which reveals within-person effects while eliminating level 2 (i.e., between-person) effects in a multilevel model (Enders & Tofighi, 2007; Nezlek 2011). As a precondition to using multilevel models, the Intraclass Correlation Coefficient (ICC) of the null model was calculated to determine the amount of variance within- and between-persons in Mood scores. The ICC was relatively low (0.28), indicating that most of the variance in Mood was due to within-person variation. Following the methodology used by Blanke et al. (2018), effect sizes were calculated via likelihood-ratio based pseudo-R^2 estimates, which approximates the unique variance accounted for by each predictor variable in the MLM. No model presented major violations of the following three MLM assumptions: linearity, homoscedasticity, and normality of residuals, predictor variables, and the dependent variable.

Results

Descriptive Data.

Of the 42 total ESM prompts sent to each participant, an average of 29.70 (70.71%) were completed (*SD* = 10.44, range = 5 - 42). The final sample consisted of 10009 total prompts. Of these 10009 total prompts, 76.40% were reported as a "Present" Attention State. This differs a bit from Killingsworth and Gilbert (2010), who reported a much lower percentage of present experiences, which we address in the *Discussion.* Within this "Present" Attention State, Inner Speech was experienced for 47.10% of the prompts. For the 23.60% of 10009 total prompts that were reported as a "Mind Wandering" attention state, Inner speech was experienced for 36.10% of the prompts. As explained above, only prompts where a participant indicated experiencing

Inner Speech (44.50% of the total 10009 prompts when collapsed across Attention State) were analyzed in the current study, and thus the current analysis was conducted on 4454 prompts.

Bivariate Associations.

As a first step, we examined bivariate associations amongst all variables collected in the ESM portion of the study (with the exception of “Current Activity”, which had to be analyzed separately, see *below*). Due to the repeated testing nature of the study, we could not rely on zero-order correlations. Instead, we used MLM’s that included one predictor variable and one dependent variable (see Table 3.1). Though we were forced to choose which variable in each pairing was considered the predictor (rows, Table 3.1) versus the dependent variable (columns, Table 3.1), the results were nearly identical if the predictor and dependent variables were swapped. For consistency, each model used a fixed slope, with participant ID entered as a random intercept effect. Note that although the reported beta coefficients are unstandardized, no data transformations were needed because all variables used the same 7-point scale, with the exception of Attention State, which was contrast coded as Present = +1, Mind Wandering = -1.

Because Current Activity was a factor with six levels, associations between this and other variables were analyzed in a different manner. *First*, to investigate the association between Current Activity and *continuous* variables, Current Activity was inputted as a fixed effect in a MLM predicting each continuous variable. A global factor-level result of Current Activity was then computed with Satterthwaite's approximation in a Type III Analysis of Variance table. *Second*, to investigate associations between Current Activity and the only other *categorical* variable (Attention State), we simply looked at whether the distribution of Activities differed substantially across the two Attention States (Table 3.2), noting that a simple statistical test (e.g., chi-square) could not be performed given the repeated data nature of the data.

The results of the associations in Table 3.1 reveal that all variables were significantly associated. Regarding hypothesis testing, we focus on three specific associations. First, in line with the hypothesis that people have improved mood when they are present, we found a significant association between Attention State and Mood, with Attention State accounting for 1.20% of the variance. Second, lending support for the hypothesis that this effect of Attention State could potentially be mediated by Thought Valence, we found 1) a significant association between Thought Valence and Mood, with more positive Inner Speech content being associated with higher Mood (Thought Valence accounting for 29.26% of the variance), and 2) a significant association between Attention State and Thought Valence, with Inner Speech being more positive for Present vs. Mind Wandering moments (Attention State accounting for 0.65% of the variance). We return to the "Mediation analysis", *below*.

Although not part of our main hypothesis testing, we also found that Clarity and Interestingness were associated with Attention State (i.e., present thoughts were clearer and more interesting). Both were also associated with Thought Valence (i.e., more positive thoughts were clearer and more interesting). Finally, Clarity and Interestingness were interrelated, i.e., clearer thoughts were more interesting. The interplay of all these additional variables in our models is described under "Covariates", below.

Table 3.1: Associations Across Key Variables

	Thought Valence	Clarity	Interestingness	Mood
Attention State	.14 [.09, .18], p < .001	.12 [.07, .18], p < .001	.09 [.04 - .14], p = .001	.21 [0.16, 0.26], p < .001
Thought Valence	-	.21 [.18, .24], p < .001	.27 [.24 - .30], p < .001	.66 [.63, .68], p < .001
Clarity	--	--	.31 [.28 - .34], p < .001	.16 [.13 - .19], p < .001
Interestingness	-	-	--	.22 [.19 - .25], p < .001

Note. Betas (unstandardized coefficients), 95% confidence intervals, and p-values are shown for predicting one variable against another in a multilevel model. The row names represent the predictor variable, and the column names represent the dependent variable in each model. For continuous variables, positive effect sizes (listed in-text above) represent that an increase in one variable was associated with an increase in the other variable (and vice versa for negative effect sizes). For Attention State, a positive effect size means that the continuous variable it is paired with is higher for Present vs. Mind Wandering moments (and vice versa for negative effect sizes). Results for Current Activity had to be analyzed differently (see Methods), and show that Current Activity was a significant predictor for each continuous variable in this Table (all *p*s < .001). The associations between different Current Activities and Attention State are presented in Table 3.2.

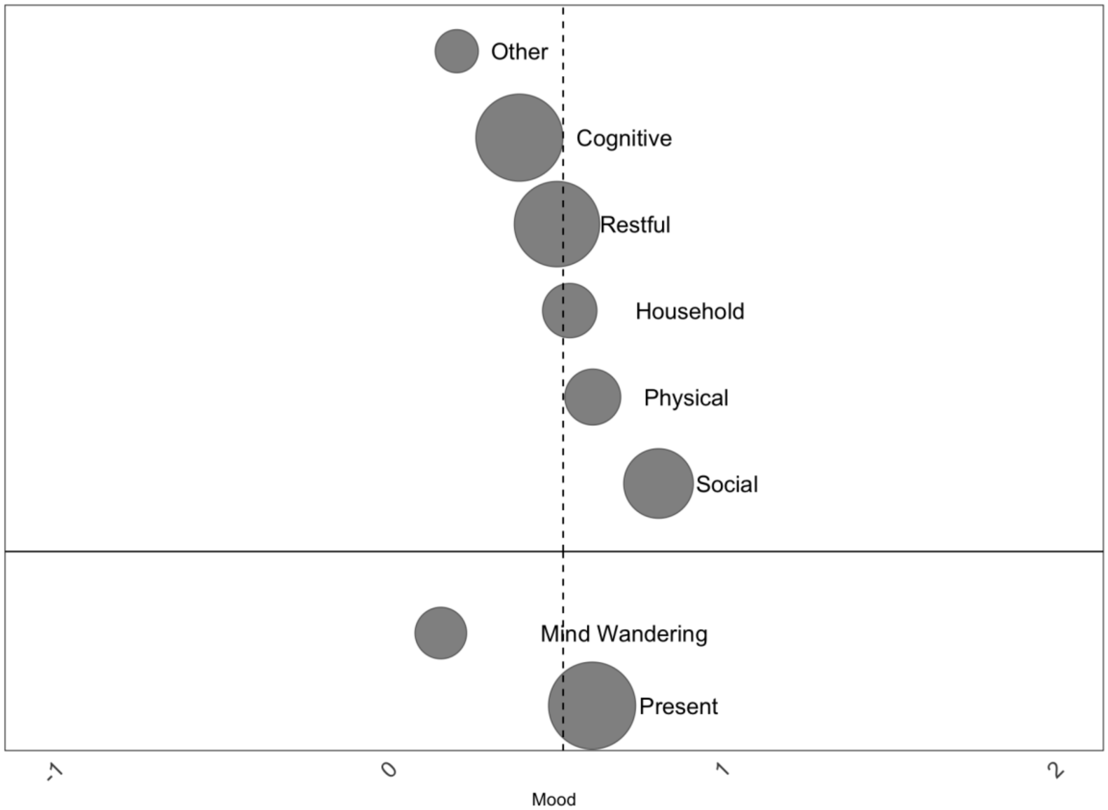
Table 3.2: Distributions of Activities across Present vs. Mind Wandering Attention States

	Mind Wandering (N=852)	Present (N=3602)	Overall (N=4454)
Activity			
Cognitive	157 (18.4%)	1221 (33.9%)	1378 (30.9%)
Household	79 (9.3%)	307 (8.5%)	386 (8.7%)
Other	43 (5.0%)	145 (4.0%)	188 (4.2%)
Physical	104 (12.2%)	315 (8.7%)	419 (9.4%)
Restful	337 (39.6%)	975 (27.1%)	1312 (29.5%)
Social	132 (15.5%)	639 (17.7%)	771 (17.3%)
Total	100%	100%	100%

Note. Being Present had a higher proportion of Cognitive Activities, whereas Mind Wandering had a higher proportion of Restful Activities. Although we could not perform chi-square statistics on these data (see Methods), this suggests an association between Current Activity and Attention State.

Visual Depictions.

In order to visualize our conceptual replication of Killingsworth and Gilbert (2010), in Figure 3.2, we plot mean Mood scores (averaged across prompts) for Present vs. Mind Wandering prompts and for each of the six Activities. Overall, the mean Mood score across all 4454 Inner Speech prompts was slightly greater than the midpoint of zero (*M* = 0.53, *SD* = 1.41), which is presented as a dashed vertical line. Collapsing across Current Activity, being Present, which accounted for 80.87% of the prompts, was associated with higher mood (N = 3602 prompts, *M* = 0.61, *SD* = 1.40) than Mind Wandering (N = 852 prompts, *M* = 0.16, *SD* = 1.40). Collapsing across Attention State, the most commonly reported Activity was “Cognitive” (30.94% of the prompts), while “Other” was the least commonly reported activity (4.22% of the prompts). “Social” Activity displayed the highest mean mood (N = 771 prompts, *M* = 0.81, *SD* = 1.55), followed by “Physical” (N = 419 prompts, *M* = 0.62, *SD* = 1.44), “Household” (N = 386 prompts, *M* = 0.55, *SD* = 1.37), “Restful” (N = 1312 prompts, *M* = 0.51, *SD* = 1.33), “Cognitive” (N = 1378 prompts, *M* = 0.40, *SD* = 1.35), and Other (N = 188 prompts, *M* = 0.21, *SD* = 1.67). Note the range of Mood scores associated with Attention State (0.45) was very similar to that for Current Activity (0.41) if “Other” is not included in the latter.



Note. Activity shown on top, Attention State on bottom. Dashed line indicates mean Mood across all prompts. Bubble size indicates the frequency of occurrence.

Figure 3.2: Mean Mood for Each Activity and Each Attention State

Covariates.

Before moving on with our main analyses, we asked whether Current Activity, Clarity, and Interestingness should be included as covariates in our multilevel models. Though we had no a priori hypotheses about these three variables, pilot data and the current data (see Table 3.1 and 3.2) show that they are associated with both Attention State and Mood. Including them in our models therefore addresses their potential confounding effects when measuring the unique contribution of Attention State to Mood. Furthermore, Current Activity was considered in

Killingsworth and Gilbert (2010). We found that including these covariates improved all model fits compared to models without the covariates. Therefore, all reported MLM analyses include these covariates with the following notable exception that we return to in the Discussion: the effect of Attention State went from predicting 1.20% of the variance in mood without including these three variables as covariates to predicting 1.07% of the variance when including them.

Attention State Analysis. Does Attention State Predict Mood?

As a first step, we ran a MLM with Mood as the dependent variable, Attention State as a contrast coded predictor variable (fixed effect), Current Activity, Clarity, and Interestingness as covariates (fixed effects), and Participant ID as a random intercept effect (Table 3.3, left panel). The results revealed a main effect of Attention State ($\beta = 0.20$, $p < 0.001$, 95% CI = [0.15 - 0.25]), which uniquely predicted 1.07% of the variance in Mood, with higher Mood for Present vs. Mind Wandering prompts. This effect, which is quite small, is markedly lower than that observed in the original Killingsworth & Gilbert (2010) study, which we address in more detail in the Discussion. Notably, all three covariates also had a significant main effect on Mood in this model, with unique variance in Mood accounted for equal to 0.8% for Current Activity, 0.6% for Clarity, and 2.1% for Interestingness.

Mediation Analysis: Does Thought Valence Mediate the Effect of Attention State on Mood?

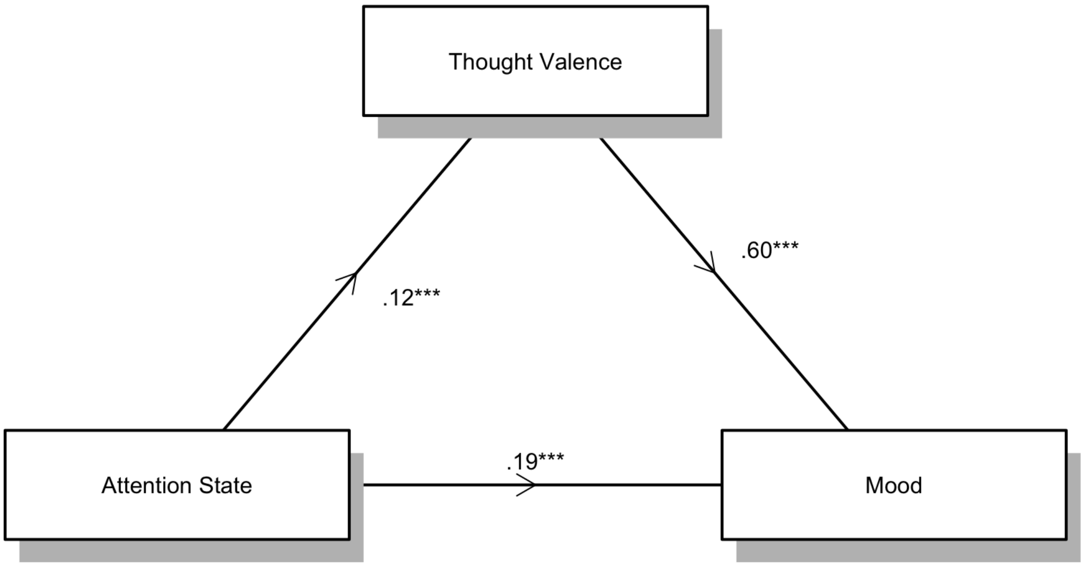
Even though the effect of Attention State on Mood was observed to be small, it was still significant, and so, we therefore moved on to ask whether this effect might be mediated by Thought Valence. To this end, in our next step we added Thought Valence to the model as a fixed effect, and the results are shown in Table 3.3 (right panel). The model revealed a main effect of Thought Valence on Mood ($\beta = 0.60$, $p < 0.001$, 95% CI = [0.57 to 0.63]), which uniquely predicted 22.16% of the variance in Mood, with higher mood for more positive inner speech content. With Thought Valence included in the model, the effect of Attention State on Mood was substantially reduced ($\beta = 0.13$, $p < 0.001$, 95% CI = 0.09 - 0.17]), uniquely predicting 0.46% of the variance. These results suggest that the relationship between Attention State and Mood may be partially mediated by Thought Valence, and that once accounted for, the unique effect of Attention State is extremely small. Of note, Thought Valence appears to have an indirect effect on the three covariates as well.

To further investigate the partial mediation of Thought Valence on Attention State, a multilevel mediation model accounting for Current Activity, Clarity, and Interestingness was analyzed using the Mediation package in R (Tingley, Dustin, et al, 2014). As Figure 3.3 illustrates, the coefficient between Attention State and Thought Valence, and the coefficient between Thought Valence and Mood, were both significant. A 95% confidence interval of the indirect effect was computed by using a quasi-Bayesian approximation using 1000 simulations. The indirect effect was (.12)*(.60) = .07, and the 95% confidence interval ranged from .05 to .11. Thus, the indirect effect was statistically significant (p <.001). That is, of the estimated 0.19 unit increase in Mood that appears to be due to Attention State, an estimated .07 of that is actually a result of Thought Valence changes generated by Attention State, while the remaining 0.12 is from Attention State itself. Put another way, the proportion of the effect of Attention State on Mood that was mediated by Thought Valence was 37.04%

Table 3.3: Variables Predicting Mood

	Attention State Model				Mediation of Valence Model			
Predictors	*Estimates*	*CI*	*p*	*Pseudo-R^2*	*Estimates*	*CI*	*p*	*Pseudo-R^2*
(Intercept)	0.13	0.02 – 0.24	**0.017**		0.33	0.23 – 0.42	**<0.001**	
Activity				0.008				0.002
Household	0.14	0.01 – 0.28	**0.042**		0.06	-0.05 – 0.17	0.305	
Physical	0.25	0.12 – 0.38	**<0.001**		0.11	0.00 – 0.22	**0.046**	
Other	-0.09	-0.28 – 0.10	0.362		-0.02	-0.18 – 0.13	0.756	
Restful	0.19	0.09 – 0.28	**<0.001**		0.12	0.04 – 0.20	**0.003**	
Social	0.33	0.22 – 0.44	**<0.001**		0.19	0.10 – 0.28	**<0.001**	
Clarity	0.09	0.06 – 0.12	**<0.001**	0.006	0.03	0.00 – 0.05	**0.037**	0.001
Interestingness	0.17	0.14 – 0.20	**<0.001**	0.021	0.07	0.05 – 0.10	**<0.001**	0.004
Attention State	0.20	0.15 – 0.25	**<0.001**	0.011	0.13	0.09 – 0.17	**<0.001**	0.005
Thought Valence					0.60	0.57 – 0.63	**<0.001**	0.222
Random Effects								
σ^2	1.31				0.91			
τ_{00}	0.52 $_{\text{Participant}}$				0.48 $_{\text{Participant}}$			
ICC	0.29				0.35			
N	337 $_{\text{Participant}}$				337 $_{\text{Participant}}$			
Observations	4454				4454			
Marginal R^2 /	0.059 / 0.328				0.274 / 0.525			

Note. Left Panel: Model asking if Attention State predicts Mood. *Right Panel*: Model asking if adding Thought Valence to the previous model lowers the effect of Attention State. Beta estimates are unstandardized. For continuous variables, positive effect sizes represent that an increase in the predictor variable is associated with an increase in Mood (and vice versa for negative effect sizes). For Attention State, a positive effect size means that Mood is higher for Present vs. Mind Wandering moments. For Activity, a positive effect size means that Activity led to higher mood than did the referent level of Cognitive Activity (and vice versa for negative effect sizes). Current Activity and Clarity are included as covariates. Bolded values are significant.



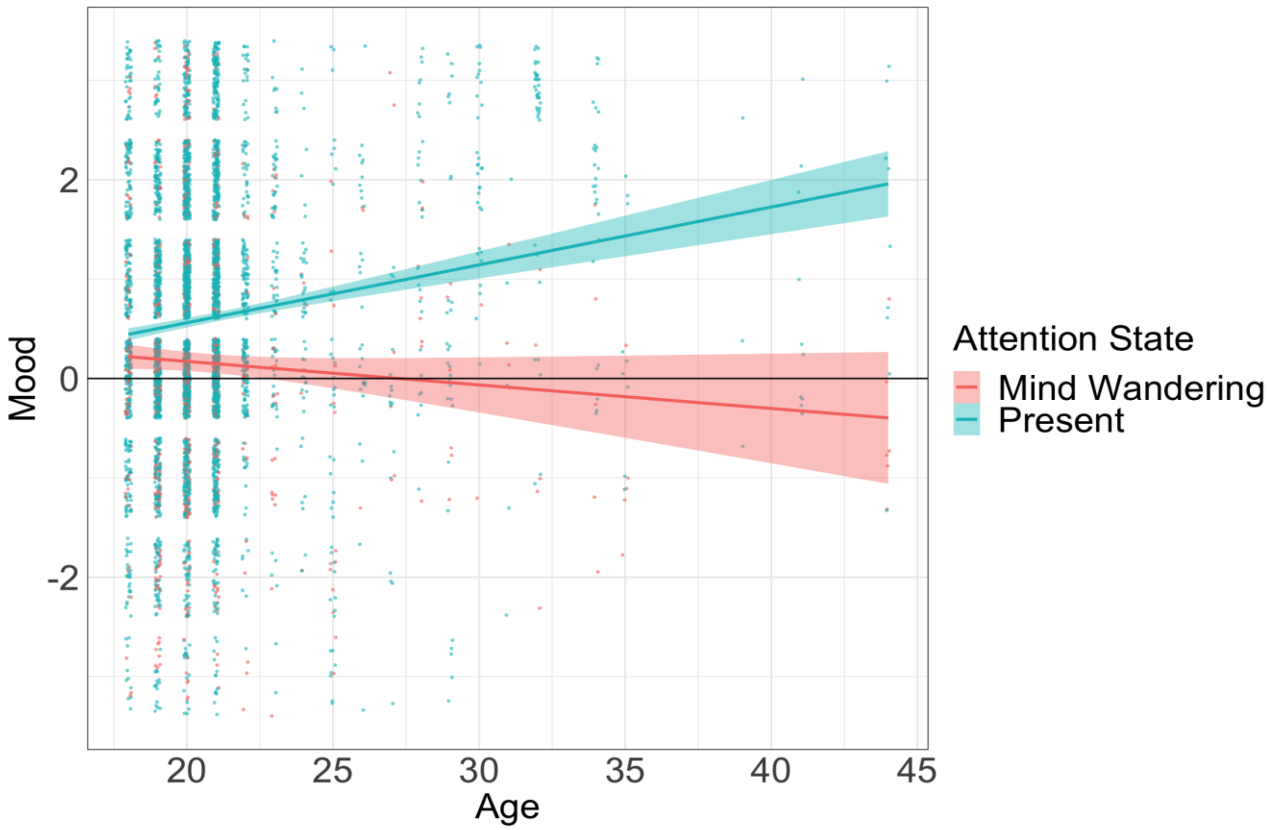
Note. The model reveals that after accounting for the effects of Current Activity, Clarity, and Interestingness, the effect of Attention State on Mood is partially mediated by Thought Valence; being present is associated with more positive thought valence, which in turn, is associated with higher mood. ***$p < .001$.

Figure 3.3: Multilevel Mediation Model

Age Moderates the Relationship between Attention State and Mood.

In exploratory analyses, we examined whether any trait-level variable (including all demographics, and the two trait measures - the FFMQ and the NEO-FFI) moderated the relationship between Attention State and Mood, with the idea that this relationship might be stronger for certain types of people. To test this, we added the trait-level variables and the cross-level interaction term for each variable with Attention State to the model in Table 3.3 (left panel) as fixed effects. The only interaction we found to be significant (with an alpha lowered to $p < .01$ given our multiple tests) was with Age ($\beta = .03$, $p < .0001$, 95% CI = [0.01, 0.04]), which uniquely predicted 0.20% of the variance in Mood. Post-hoc analysis revealed two ways to describe the interaction (presented in Figure 3.4). The first way is to note that, for Present moments, older people are happier than younger people (and the reverse age trend is seen for

Mind Wandering moments). The more obvious way to explain the interaction is to note that the relationship between Attention State and Mood grew stronger with increasing age (this effect was seen despite the age range being relatively limited given the student sample used). This effect of age, together with the fact that our sample was much younger than that of Killingsworth and Gilbert (2010), might explain why they found that Attention State explained more variance in Mood than found in the current study, which we address in the *Discussion.*



Note. The effect of Attention gets stronger with age.

Figure 3.4: Age moderates the relationship between Attention State and Mood

Discussion

Although several studies have examined the effects of attention state on mood, the present research represents one of the first attempts to investigate how various thought qualities that naturally occur across different attention states impact concurrent mood. The results of the current ESM study corroborate those of Killingsworth and Gilbert (2010) showing that attention state does matter; people are less happy when mind wandering than when present. Unlike their study, however, which asked participants about thought valence only for mind wandering moments, we inquired about thought valence for both mind wandering and present moments. This allowed us to reveal that the relationship between attention state and concurrent mood is partially accounted for by thought valence. More precisely, mind wandering appears to be associated with more negative thoughts, which helps explain its association with worse mood.

While our findings demonstrate that the inclusion of thought valence substantially reduces the strength of the relationship between attention state and concurrent mood, we would like to emphasize that the current study was not designed to speak to the question of causality. While our correlational findings may suggest that mind wandering leads to more negative thoughts, in turn driving a worse mood, it is also possible that negative thoughts lead the mind to wander, in turn driving a worse mood. For example, if a troubling thought arises while an individual is attempting to complete a task, this may create a particularly compelling urge to prioritize thinking through the concern at the expense of paying attention to the task at hand. To establish causality, future experimental studies must show either that 1) inducing a change in one variable creates a change within another or 2) that a state of one variable at time 0 predicts the state of another at time 1.

Several studies have used the latter method, referred to as time-lagged designs, to examine the various parts of our proposed meditation model. Using experience-sampling in everyday life, some studies have reported a relationship between *attentional state* and *subsequent mood* (Welz et al., 2018), while others have failed to observe this relationship (Mills et al., 2021; Poeri et al., 2013). Poeri et al. (2013) did, however, find evidence for the opposite causal direction, i.e., poor mood predicting later occurrences of mind wandering, but others have also failed to observe this relationship (Killingsworth & Gilbert, 2010; Mills et al., 2021). Such mixed results may be due to differences in the amount of time between prompts (which should be fairly immediate, see Mason et al., 2013 for more discussion on this criticism), due to wide variation in activities one may be engaged in during normal life, or due to whether concurrent mood was controlled for or not.

To offer more control, causal designs are often carried out in tightly-controlled laboratory-based contexts. In these studies, participants are randomly prompted while they engage in a cognitive task that is minimally demanding, and thus likely to induce mind wandering. Such studies have shown a more consistent negative relationship between the attentional state and subsequent mood (Ruby et al., 2013; Marchetti et al. 2012; Wilson et al. 2014), with greater task-unrelatedness predicting poorer mood (while Ruby and colleagues further found that poorer mood predicted later levels of task-unrelated thought). However, several criticisms have arisen regarding the generalizability of findings regarding task-unrelated thoughts induced in a laboratory setting compared to other forms of naturally occurring mind wandering (e.g., see Murray et al., 2020). Future studies should be wary of addressing these various limitations.

In terms of the causal association between *attention state* and *thought valence*, no ESM studies to date have examined the time-lagged effect of attention on later thought valence or vice versa. While one can imagine that mind wandering may *lead* to negative thoughts, it may also be the case that having negative thoughts leads to mind wandering. This direction of causality is supported by studies showing that inducing negative thought valences in people (e.g., by delivering distressing news, presumably resulting in negative thoughts) leads to greater mind wandering during task engagement (e.g., Antrobus et al., 1966; Smallwood et al., 2009).

Several laboratory-based studies have used time-lagged designs to examine *thought valence* as a predictor of *subsequent mood*. For example, Marchetti et al. (2012) reported a significant association between thought valence and subsequent positive, but not negative, mood states, while Ruby et al. (2013) similarly found that more positive (pleasant) thoughts predict more subsequent positive mood ratings. In contrast, in a recent meta-analytic review examining the pooled correlation between valence and mood, negatively valenced mind wandering (referred to as “unprompted thought” in this paper) had a bigger impact on mood than positively valenced mind wandering (Kam et al., 2024). Furthermore, using experience sampling, Poeri et al. (2013) found that negatively valenced mind wandering predicts a subsequent detriment to mood, while also finding that a sad mood precedes negatively valenced mind wandering (i.e., mind wandering to current concerns); in other words, suggestive evidence for both causal directions between thought valence and mood have been found. However, it is important to note that both in Poeri et al. (2013) and in the review by Kam et al. (2024), the analyses only included data from *mind wandering states*, similar to Killingsworth & Gilbert. Future research is necessary to replicate these results in ecologically valid contexts, while also exploring how these relationships change across attentional states.

Regardless of the direction of causality, the prominent role of thought valence in predicting mood suggests that rather than changing the frequency of a particular attention state (such as through mindfulness practices aimed at promoting increased frequency of present focused attention), individuals may want to focus more on changing the quality of one’s thoughts - a notion that is in line with cognitive behavioral therapy (CBT; Beck, 1997; Hoffmann et al., 2012). At first glance, this suggestion might seem contradictory to a large volume of research showing the benefits of mindful meditation on improving mood (Eberth & Sedlmeier, 2012; Pascoe et al., 2021; Rodrigues, Nardi, & Levitan, 2017). However, mindfulness is multifaceted with most common definitions consisting of at least two facets: present-focused attention *and* an attitude of acceptance/nonjudgement (Bishop 2004). Future research should continue to explore the relative importance of each component when measured in naturalistic settings using state, rather than trait, measures of mindfulness. Regardless of their relative importance, it is promising that with training, individuals have the ability to enhance both their capacity for present focused attention (e.g., through mindfulness training), and for amplifying positive inner speech patterns (e.g., through CBT).

More broadly, our findings suggest that the commonly reported detrimental impact of mind wandering on concurrent mood may largely be accounted for by certain confounding variables. Indeed, the present research demonstrated that in addition to thought valence playing a substantial role in explaining the relationship between attention state and concurrent mood, thought interestingness and clarity also showed meaningful effects, with more clear and interesting thoughts associated with better mood. Similarly, Franklin et al. (2013) found that off-task thoughts rated as more interesting during daily life were associated with a better mood (however no previous studies have examined these qualities across both mind wandering and present focused attention states). Such findings raise intriguing possibilities for future research, such as understanding the circumstances and conditions under which mind wandering can provide positive benefits. Despite the generally poor reputation surrounding mind wandering, several studies indicate that mind wandering can offer benefits such as entertainment (Franklin et al., 2014), feelings of social bond and connection (Poerio et al., 2015), and emotional respite from boring or stressful circumstances (e.g. Molstad, 1986), as well as being potentially important for creative thinking (e.g. Baird et al., 2012; but see Murray et al., 2021 for null results). These findings are underscored by recent studies indicating that the beneficial side of mind wandering can even be enhanced. For instance, a recent study showed that certain circumstances, such as listening to positive music, can promote more pleasant mind wandering (Koelsch et al., 2019; Taruffi, 2021; Taruffi et al., 2017). Future correlational research is required to demonstrate how such unmeasured third variables may further reduce the remaining unexplained variance in mood using naturalistic data collection modalities.

That said, although the inclusion of additional variables in the present study substantially reduced the magnitude of association between attention state and mood, the main effect of attention state remained small but significant. This may suggest that there is an intrinsic effect of attentional state on wellbeing, as Killingsworth & Gilbert originally claimed, but the magnitude of this effect on mood seems to be greatly overstated in the literature. Furthermore, it is important to note that the overall effect of attention state on mood was much smaller in the current study than in Killingsworth & Gilbert. While Killingsworth & Gilbert (2010) did not report on the Pseudo-R^2 statistic reported in our results, through personal communication, they re-ran these statistics using their original data and models. They found that a binary mind-wandering variable uniquely predicted 4% of the variance in mood, and the valenced mind-wandering variable uniquely predicted 13% of the variance in mood. One possible explanation for this discrepancy in effect sizes is that the effect of attention on mood may be contingent on sample demographics. The current study explored this prospect by examining a number of moderating trait variables. Of all those explored, the only one that showed a marginal effect was age, whereby older participants showed the largest effects of attention state on mood. This effect of age, together with the fact that the mean age of our sample (21 years) was much younger than the mean age of the Killingsworth and Gilbert sample (34 years), might partially explain why attention state explained more variance in mood in their study compared to the current study.

In addition to differences in effect sizes between Killingsworth and Gilbert (2010) and the current study, we also observed a lower frequency of mind wandering prompts (23.6% of all prompts; 36.1% of Inner Speech prompts) than they did (46.9% of prompts). Several other studies have also reported much lower rates of mind wandering to Killingsworth & Gilbert (2010; e.g., Kane et al., 2007). While some of this difference might be explained by methodological or age differences, there is also some evidence that this discrepancy may be due to our sample demographic. For example, in an ESM study with a similar demographic makeup to ours (undergraduate Chinese students, 70% female), an average of 24.4% of prompts involved mind wandering (Song & Wang, 2012). Further studies are required to understand who, when, and why some people mind wander more frequently than others.

Finally, it is important to note that how mood is operationalized varies across studies. Affective theories commonly distinguish at least two components of mood: arousal and valence (e.g., Posner et al., 2005; Russell, 1980). We opted for the more commonly used dimension of affective valence of mood. This choice was guided by previous research demonstrating inconsistent relationships between attention state and arousal (e.g., Mills et al., 2021; Franklin, Broadway, et al., 2013, Mittner et al., 2014, and Unsworth & Robison, 2018). Nevertheless, variation in measurement choices could explain differing results observed across studies, suggesting that this is an area worth further consideration.

In conclusion, the present findings suggest that although a wandering mind is not necessarily an unhappy mind, when the stream of consciousness diverges from the present it is somewhat more likely to flow to unhappy places. In large part, the reduced mood associated with mind wandering appears tied to the negative thought qualities that frequently accompany such departures from the here and now. However, this is clearly not the whole story as mind wandering was still associated (albeit to a markedly lesser degree) with reduced mood even when thought quality and current activity (as well as clarity and interestingness) was controlled for. It remains uncertain why, all other things being equal, a distracted mind tends to be a bit less happy. However, what is clear is that wandering minds are markedly happier when entertaining more positive material. It therefore seems prudent that we all find cheerful things to mind wander about.

Acknowledgements

Chapter 3 is currently under revision and is anticipated to appear in Emotion, 2024, Raynes, Stephen; Gross, Madeleine; Schooler, Jonathan; Guo, Evie; Dobkins, Karen, APA PsycArticles. The dissertation author was the primary researcher and author of this papers.

References

Antrobus, J. S. (1968). Information theory and stimulus-independent thought. British Journal of Psychology, 59(4), 423-430.

Antrobus, J. S., Singer, J. L., & Greenberg, S. (1966). Studies in the stream of consciousness: Experimental enhancement and suppression of spontaneous cognitive processes. *Perceptual and Motor Skills*, *23*(2), 399-417.

Baer, R. A., Carmody, J., & Hunsinger, M. (2012). Weekly change in mindfulness and perceived stress in a mindfulness-based stress reduction program. Journal of Clinical Psychology, 68(7), 755–765. https://doi.org/10.1002/jclp.21865

Baird, B., Smallwood, J., Mrazek, M. D., Kam, J. W., Franklin, M. S., & Schooler, J. W. (2012). Inspired by distraction: Mind wandering facilitates creative incubation. *Psychological science*, *23*(10), 1117-1122.

Bates, D., Maechler, M., Bolker, B., Walker, S., Christensen, R. H. B., Singmann, H., & Green, P. (2009). Package 'lme4'. *URL http://lme4. r-forge. r-project. Org*.

Beck A. T. (1997). The past and future of cognitive therapy. *The Journal of psychotherapy practice and research*, *6(4)*, 276–284.

Bishop, S. R., Lau, M., Shapiro, S., Carlson, L., Anderson, N. D., Carmody, J., ... & Devins, G. (2004). Mindfulness: A proposed operational definition. *Clinical psychology: Science and practice*, *11*(3), 230.

Blanke, E. S., Riediger, M., & Brose, A. (2018). Pathways to happiness are multidirectional: Associations between state mindfulness and everyday affective experience. *Emotion*, *18*(2), 202.

Costa, P. T., & McCrae, R. R. (1992). Revised NEO Personality Inventory (NEO-PI-R) and NEO Five-Factor Inventory (NEO-FFI) professional manual. Odessa, FL: Psychological Assessment Resources.

Eberth, J., & Sedlmeier, P. (2012). The effects of mindfulness meditation: a meta-analysis. *Mindfulness*, 3(3), 174-189.

Enders, C. K., & Tofighi, D. (2007). Centering predictor variables in cross-sectional multilevel models: a new look at an old issue. *Psychological methods*, *12*(2), 121.

Fox, K. C., Thompson, E., Andrews-Hanna, J. R., & Christoff, K. (2014). Is thinking really aversive? A commentary on Wilson et al.'s "Just think: the challenges of the disengaged mind". *Frontiers in psychology*, *5*, 1427.

Franklin, M. S., Mrazek, M. D., Anderson, C. L., Smallwood, J., Kingstone, A., & Schooler, J.

W. (2013). The silver lining of a mind in the clouds: Interesting musings are associated with positive mood while mind-wandering. *Frontiers in psychology, 4*, 583

Giambra, L. M. (1989). Task-unrelated thought frequency as a function of age: a laboratory study. Psychology and aging, 4(2), 136.

Hobbiss, M. H., Fairnie, J., Jafari, K., & Lavie, N. (2019). Attention, mindwandering, and mood. *Consciousness and cognition, 72*, 1-18.

Hofmann, S. G., Asnaani, A., Vonk, I. J., Sawyer, A. T., & Fang, A. (2012). The Efficacy of Cognitive Behavioral Therapy: A Review of Meta-analyses. *Cognitive therapy and research, 36(5),* 427–440. https://doi.org/10.1007/s10608-012-9476-1

Kam, J. W., Wong, A. Y., Thiemann, R. F., Hasan, F., Andrews-Hanna, J. R., & Mills, C. (2024). On the relationship between unprompted thought and affective well-being: A systematic review and meta-analysis. *Psychological Bulletin*.

Kane, M. J., Brown, L. H., McVay, J. C., Silvia, P. J., Myin-Germeys, I., & Kwapil, T. R. (2007). For whom the mind wanders, and when: An experience-sampling study of working memory and executive control in daily life. *Psychological science*, *18*(7), 614-621.

Killingsworth, M. A., & Gilbert, D. T. (2010). A wandering mind is an unhappy mind. *Science*, *330*(6006), 932-932.

Klinger, E. (1977). *Meaning and void: Inner experience and the incentives in peoples lives*. U of Minnesota Press.

Koelsch, S., Bashevkin, T., Kristensen, J., Tvedt, J., & Jentschke, S. (2019). Heroic music stimulates empowering thoughts during mind-wandering. *Scientific Reports*, *9*(1), 1-10

Marchetti, I., Koster, E. H., & De Raedt, R. (2012). Mindwandering heightens the accessibility of negative relative to positive thought. *Consciousness and cognition*, *21*(3), 1517-1525.

Mason, M. F., Bar, M., & Macrae, C. N. (2009). Exploring the past and impending future in the here and now: Mind-wandering in the default state. *Cognitive Science Compendium*, *2*, 143-162.

Mason, M. F., Brown, K., Mar, R. A., & Smallwood, J. (2013). Driver of discontent or escape vehicle: the affective consequences of mindwandering. *Frontiers in psychology*, *4*, 477.

Molstad, C. (1986). Choosing and coping with boring work. J. Contemp. Ethnogr. 15, 215–236. doi: 10.1177/089124168601500204

Mooneyham, B. W., & Schooler, J. W. (2013). The costs and benefits of mind-wandering: a review. Canadian Journal of Experimental Psychology/Revue canadienne de psychologie expérimentale, 67(1), 11.

Murray, S., Liang, N., Brosowsky, N., & Seli, P. (2021). What are the benefits of mind wandering to creativity? Psychology of Creativity, Esthetics, and the Arts. Advance online publication

Nezlek, J. B. (2011). *Multilevel modeling for social and personality psychology*. SAGE Publications Ltd.

Oerlemans, W. G., Bakker, A. B., & Veenhoven, R. (2011). Finding the key to happy aging: A day reconstruction study of happiness. *Journals of Gerontology Series B: Psychological Sciences and Social Sciences, 66*(6), 665-674.

Pascoe, M. C., de Manincor, M., Tseberja, J., Hallgren, M., Baldwin, P. A., & Parker, A. G. (2021). Psychobiological mechanisms underlying the mood benefits of meditation: a narrative review. *Comprehensive Psychoneuroendocrinology, 6*, 100037.

Poerio, G. L., Totterdell, P., Emerson, L. M., & Miles, E. (2015). Love is the triumph of the imagination: Daydreams about significant others are associated with increased happiness, love and connection. *Consciousness and Cognition*, *33*, 135-144.

Rodrigues, M. F., Nardi, A. E., & Levitan, M. (2017). Mindfulness in mood and anxiety disorders: a review of the literature. *Trends in psychiatry and psychotherapy, 39,* 207-215.

Ruby, F. J., Smallwood, J., Engen, H., & Singer, T. (2013). How self-generated thought shapes mood—the relation between mind-wandering and mood depends on the socio-temporal content of thoughts. *PloS one, 8*(10), e77554.

Smallwood, J., Fitzgerald, A., Miles, L. K., & Phillips, L. H. (2009). Shifting moods, wandering minds: negative moods lead the mind to wander. *Emotion*, *9*(2), 271.

Taruffi, L. (2021). Mind-wandering during personal music listening in everyday life: Music-evoked emotions predict thought valence. *International journal of environmental research and public health, 18*(23), 12321.

Taruffi, L., Pehrs, C., Skouras, S., & Koelsch, S. (2017). Effects of sad and happy music on mind-wandering and the default mode network. *Scientific reports*, *7*(1), 14396.

Tingley, D., Yamamoto, T., Hirose, K., Keele, L., & Imai, K. (2014). Mediation: R package for causal mediation analysis.

Welhaf, M. S., Banks, J. B., & Bugg, J. M. (2024). Age-Related Differences in Mind Wandering: The Role of Emotional Valence. *The Journals of Gerontology: Series B*, *79*(1), gbad151.

Wilson, T. D., Reinhard, D. A., Westgate, E. C., Gilbert, D. T., Ellerbeck, N., Hahn, C., ... & Shaked, A. (2014). Just think: The challenges of the disengaged mind. *Science*, *345*(6192), 75-77.

CONCLUSION

In sum, this dissertation (1) provides evidence that a newly created state mindfulness measure is psychometrically sound; (2) clarifies the relative importance of each facet of state mindfulness in predicting state affect within experimental and naturalistic settings; and (3) determines how much of the relationship between state mindfulness and state affect can be accounted for by valence.