



Daily Mindfulness Components and State Affect: A Day Reconstruction Study

Stephen Raynes¹ · Karen Dobkins¹

Accepted: 10 February 2025
© The Author(s) 2025

Abstract

Objectives The assessment of state mindfulness often involves detecting mindful experiences during formal practices such as 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.

Method On 2 consecutive days, undergraduate students were asked to recall episodes from the previous day, and for each episode, report on their level of state mindfulness and state affect, as well as other covariates likely to predict affect. State mindfulness was assessed with the newly created State Four Facet Mindfulness Questionnaire (state-4FMQ), adapted from the commonly used trait Five Facet Mindfulness Questionnaire. Before and after completing the 2-day DRM task, participants completed several trait measures.

Results Multilevel modeling revealed that daily experiences of state ActAware and Nonjudgment, and to a lesser degree Observing, uniquely predicted state affect. Describing was not uniquely predictive of state affect. Various analyses on the current data also provided further validation of the state-4FMQ. Analysis of variance (ANOVA) tests revealed there was no significant long-term benefit to completing the DRM per se.

Conclusions 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-4FMQ within the DRM.

Preregistration This study was preregistered at the Open Science Framework: <https://osf.io/j7fn>.

Keywords Mindfulness · State Four Facet Mindfulness Questionnaire (state-4FMQ) · Day reconstruction method (DRM) · State affect

A critical question in mindfulness research is whether people are happier when they are more mindful. Research has generally explored this question within one of two contexts. The first context is “formal mindfulness,” which refers to 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

refers to 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 widely accepted definition 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

✉ Karen Dobkins
kdobkins@ucsd.edu

¹ Department of Psychology, University of California, San Diego, 9500 Gilman Dr, La Jolla, CA 92093, USA

rigorous and ecologically valid way, state measures of daily mindfulness are preferable.

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 moment-to-moment fluctuations in daily mindfulness (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 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). Note that a third method, daily end-of-day diary (EOD; also called the daily diary method), is 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 reinstatement (DRM) or momentary recall (ESM), thereby making it the most prone to bias.

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). ESM is typically considered the gold standard among ILD methods because it can capture an experience in situ, be relatively easy for participants to complete, and be coupled with other real-time ambulatory measures (e.g., heart rate variability). The question of whether people are happier when they are more mindful has been amply explored with ESM designs (e.g., Gross et al., 2024; 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 state mindfulness and well-being (including state affect) with wide-ranging effect sizes depending on several aspects of the study design. Although these effects are promising, one gap in this literature is the inconsistency in how state mindfulness has been operationalized, including whether or not state mindfulness is treated as a

multi-dimensional construct. Here, we argue that state mindfulness is likely to be multidimensional (as is claimed to 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 is likely attributed to the lack of a validated, multidimensional state mindfulness measure suitable for ILDs without imposing excessive participant burden.

In ESM, although individual prompts are typically short, these studies are often described as time-consuming (Kahneman et al., 2004) and burdensome to participants, and therefore have the potential for high attrition. Participants can also be burdened by the intrusiveness of needing to carry a device that interrupts their day at unpredictable times to complete the same prompted survey (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 (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 limiting the collection of precise time-use data (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.

In contrast, the DRM was developed as a means of reproducing the information that would be collected through ESM, but without the shortcomings described above (Kahneman et al., 2004). Based on techniques grounded in cognitive science (see commentary by Diener & Tay, 2014; 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 1 day closely correspond to data collected from separate studies, and separate samples, that use 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 between the two methods (Bylsma et al., 2011; Dockray et al., 2010; Kim et al., 2013; Lucas et al., 2021; Schneider et al., 2020). Notably, the agreement of affect ratings between ESM and DRM methods differs depending on which aspect of state affect is assessed; for example, Dockray et al. (2010) found that happiness, yet not anger, is nearly 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.

We next highlight the many practical (and empirical) benefits to using DRM that are not easily afforded in ESM. 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 (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, 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 (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 first goal of the current study was to investigate, using the DRM, whether people are happier in daily life episodes 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-4FMQ”; Raynes & Dobkins, 2025), 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 among a general population. Both the state-4FMQ and trait-FFMQ include the following four facets: *Acting with Awareness (ActAware)* 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. The trait-FFMQ has a fifth factor, *Non-reactivity*, which is the ability to allow thoughts and emotions to come and go without becoming attached or carried away with them. Through exploratory factor analysis, this facet was omitted in the state-4FMQ. In the development article of the state-4FMQ (Raynes & Dobkins, 2025), state affect following a single meditation session was strongly and equally predicted by ActAware and Nonjudgment, 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-4FMQ facet in the context of formal mindfulness. The current study asked this same question but in the context of daily mindfulness, with the expectation of finding the same pattern of results.

This second goal of the current study was to provide further validation of the newly created state-4FMQ (Raynes & Dobkins, 2025). Here, we tested the predictive, incremental, convergent, and construct validity of the state-4FMQ 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 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.

To summarize, there were three main goals in this study: first, to estimate the differential associations between state mindfulness facets and state affect, in particular happiness; second, to further validate the state-4FMQ as a multidimensional measure of state mindfulness; and third, to investigate whether the DRM itself has positive long-term effects on participants.

Method

Participants

Participants were undergraduate students recruited in 2023–2024 through the UCSD participant pool, an online tool run by the Department of Psychology where undergraduate students sign-up to participate in research studies in exchange for course credit. Eligibility was restricted to participants who reported being at least 18 years old. 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 the approach used by Blanke et al. (2018) rather than using a formal power analysis, because the latter is complex and controversial for multi-level models (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. Although our pre-registration estimated a slightly larger sample based on conservative assumptions, we revised our recruitment strategy prior to data collection in light of updated methodological considerations, yielding a final sample that—though somewhat lower than originally planned—remained sufficient for our analyses.

The following five exclusion criteria (as outlined in our preregistration) 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, 2025). 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 UCSD subject pool 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 studies (and were unable to sign up for more than one study), which differentiated their condition. All three studies were posted simultaneously (which randomizes the order of available studies each login), listed as a 5-day experiment with the same overall time commitment, were worth the same amount of course credit, had the same sign-up timeframe, had the same study abstract, and had as similar a study description as possible. 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 5 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 address the three goals of the current study. Our *first* goal was addressed with the data from Days 2 and 3 collected from participants in the “DRM condition” (see below), allowing us to investigate whether

the different components of state mindfulness uniquely predict state affect, specifically happiness, in daily life. Our *second* goal was addressed with the data from Days 1, 2, and 3 from participants in the “DRM condition,” allowing us to conduct further validation analyses of the state-4FMQ. Our *third* goal was addressed with the data from Days 1, 4, and 11, allowing 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 2 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 15 min and 2 hr, 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 the time of the episode, how much they remembered of it, and the positive versus negative valence of the activity they were involved in (see “Measures,” below). If a participant was awake for over 24 hr, they were asked to enter episodes for the first 24 hr they were awake. This portion of listing out the episodes is referred to as the “reinstatement task.”

After the reinstatement, they then completed what we refer to as the “experience reporting task,” performed separately and chronologically 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. Note that, after completing the reinstatement task, participants in the DRM group were instructed that they would be asked questions about a random drawing of listed episodes, although, in reality, all episodes that passed this integrity check were included. We informed participants of this in a way that maintained the integrity of the study design because we feared that if they figured out on the first DRM day that there were not asked to perform the experience reporting task on episodes they could not remember well, that they would falsely state they could not remember episodes on the second DRM day to get out of having to do the next task. For each sufficiently remembered episode, the experience reporting task consisted of participants answering a single question about state affect, 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 “Active Control” and the “Passive Control” conditions. In the Active Control condition, on both Days 2 and 3, participants *only* completed the reinstatement task described above, and not the experience reporting task. This 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 Active Control condition). In the Passive Control condition, participants did not complete any task on Days 2 and 3. This 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 2 days.

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 Passive 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.

Measure

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 to 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 scores ranging from $\alpha = 0.77$ to 0.84, and $\omega = 0.83$ to 0.89, for the total across the three timepoints in the current 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 to 7. Responses are averaged to provide a single total score, with higher scores reflecting greater overall happiness. The scale showed acceptable internal consistency scores ranging from $\alpha = 0.85$ to 0.87, and $\omega = 0.87$ to 0.88, across the three timepoints in the current study.

Reinstatement Task Measures

For the reinstatement task (which both DRM and Active Control participants completed), for each episode, participants were asked to label the episode with an open-end text response, and select the approximate start and end time with dropdown selections in 5-min increments. They were then asked two questions. First, “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. Second, they were asked about the *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.

Experience 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). Despite the inherent shortcomings of measuring any construct via a single-item such as reduced reliability and validity, 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 of the State Four Facet Mindfulness Questionnaire (state-4FMQ; Raynes & Dobkins, 2025), which is a state adaptation of the Five Factor Mindfulness Questionnaire (trait-FFMQ; Baer et al., 2006), and employs four of the five facets (Act-Aware; Describing; Nonjudging; Observing; and omitting Nonreactivity) 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 was used in a recent ESM study (Gross et al., 2024): (1) *Social activity*, (e.g., *being around family, friends, or peers; 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).

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 et al. (2010), and Pang and 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; Raynes & Dobkins, 2025; 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.” 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-4FMQ 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 Analyses

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% ($p < 0.05$) for all two-tailed tests; however, we emphasize the effect sizes rather than statistical significance since the latter is often misleading. Effect sizes (rules of thumb) are as follows: For Pearson’s R values, 0.10–0.30 are weak effects, 0.30–0.50 are medium effects, and 0.50 and over are large effects (Cohen, 1988). For Cramer’s V for chi-square tests, values ≥ 0.1 are weak, ≥ 0.3 are moderate, and ≥ 0.5 are large effects (Kakudji et al., 2020). For partial eta squared (η^2) for analysis of variance (ANOVA) tests, 0.01 indicates a small effect, 0.06 indicates a medium effect, and 0.14 indicates a large effect (Cohen, 1988).

All analyses were computed using R (Version 4.2.2; R Core Team, 2022). Multilevel models (MLM’s) were used when considering within-person analyses to account for the natural two-level data structure, where prompts collected over time are nested within individuals, and includes the following analyses: All of Goal 1 (Do Different Components of State Mindfulness Uniquely Predict State Affect?), and some of Goal 2 (Further Validation of the state-4FMQ; specifically, predictive validity, and tests 1–2 of construct validity). All MLM’s were assessed using the R-package *lme4* (v1.1–27.1; Bates et al., 2018) 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 & Tofghi, 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 (Hudson et al., 2020).

For consistency and ease of interpretability and reproducibility, all MLM's used fixed slopes, with participant ID entered as a random intercept effect. This statistical decision was largely informed by our research being in an early stage of exploration, both in terms of using a relatively rarely used methodology (the DRM), and a new psychometric scale (the state-4FMQ). Although the inclusion of random slopes may provide a more nuanced understanding of how individual differences may interact with the predictor variables in predicting state affect, they also significantly increase model complexity and thus risk overfitting and jeopardizing the robustness of fixed effect estimates, thereby inhibiting the replicability of our results. Furthermore, more complex multilevel models, such as those incorporating random slopes, can affect type I error rates and power in nuanced ways (e.g., Barr et al., 2013; Bates et al., 2018; Matuschek et al., 2017). Once a more foundational understanding of the state-4FMQ is established, future researchers should compare models with different complexities of random effects. 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 MLM's 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 MLM's, with the expectation of detecting varying effect sizes across the facets. We expected a positive relationship between the state-4FMQ and state affect.

The resilience of these results was tested by assessing whether the state-4FMQ facet scores still uniquely predicted variance in state affect after accounting for several relevant fixed effect covariates.

The first covariate was Trait Mindfulness (trait-FFMQ, Baer et al., 2006, obtained on Day 1), which was included in an "incremental" model (see below) as a means of disentangling the effects of state versus trait aspects of mindfulness, which is specifically relevant to testing the "Construct Validity" of the newly created state-4FMQ (below).

The other covariates were (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., 2024; 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 (both the activity itself and the valence of that activity) in that moment. With that in mind, the purpose of including these covariates was to minimize unaccounted variance in state affect scores, so that we could more clearly measure the unique effects of the state mindfulness facets on state affect.

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-4FMQ total score in predicting state affect; if it did interact, we would learn that the covariate is acting as a moderator of the relationship between state mindfulness and state affect. 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 Type/Valence, entered simultaneously).

Further Validation of the state-4FMQ

Because the state-4FMQ 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-4FMQ.

Predictive Validity Finding evidence that several components of state mindfulness uniquely predict state affect provides confirmation for the predictive validity of the state-4FMQ, as was observed in the development of the state-4FMQ (Raynes & Dobkins, 2025).

Construct Validity Construct validity of the state-4FMQ was examined in three ways. First, the multidimensionality of the state-4FMQ 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., a MLM of the total state-4FMQ score). We then quantified the difference in model fit by computing the differences of respective AIC and BIC values. 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. Note that a more conventional approach to testing multidimensionality via EFA and CFA was investigated in the development article of the state-4FMQ (Raynes & Dobkins, 2025).

Second, the state- vs. trait-like behavior of the state-4FMQ was tested in an incremental model, asking whether the state-4FMQ predicts state affect over and beyond that predicted by the trait-FFMQ. If so, this would provide evidence that the state-4FMQ is not masquerading as a trait measure, as was found in the development article of the state-4FMQ (Raynes & Dobkins, 2025). Evidence for state-like behavior was also assessed by testing whether each component of the state-4FMQ 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 ($\sim < 0.50$) 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-4FMQ to detecting mindful states. Inspired by Burzler and Tran (2022), the mean state-4FMQ 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-4FMQ 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-4FMQ by asking whether each state-4FMQ facet aligned well with the same facet from the trait-FFMQ facet. For this, we calculated bivariate correlations between the Level 2 state-4FMQ 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 medium to large positive correlation among each aligned facet (e.g., between mean state Observing vs. trait Observing). 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, 2002a, 2002b).

We also separately explored the convergent validity of our in-house state affect measure using this same methodology, with the expected result of a medium to large 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, Active Control and Passive 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 2-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, Active Control, Passive 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.

The study design, hypotheses, and analysis plan were prospectively preregistered (prior to data collection) and are publicly available here: <https://osf.io/t6x4h/>.

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 Passive control (DRM $n = 113$; Active Control $n = 107$; Passive Control $n = 150$) was unexpected. Although the quasi-randomization procedure did not unfold as planned, the sample sizes for each group were still large enough to conduct all planned analyses with integrity; hence, we proceeded without modification. Demographic information can be found in Table 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 min. To ensure participants sufficiently remember the episode they were reporting on, the reinstatement 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 episodes, 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%).

In the Active 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 Active Control day during the Intervention Protocol was 11.11 min.

As an exploratory measure, we asked participants from all three groups on Day 4 about the typicality of the past 2 days. Overall, while most participants ($n = 298$, 80.50%) reported that the past 2 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 2 days. A chi-square test revealed no significant differences in these proportions between the three participant groups, $\chi^2(4, n = 370) = 2.78$, $p = 0.60$.

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 the state-4FMQ

Table 1 Demographic information for the three conditions

	DRM ($n = 113$)	Active control ($n = 107$)	Passive 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%)

on state affect in a MLM. For this MLM, state affect was the dependent variable, the four state-4FMQ facets were entered simultaneously as predictor variables (fixed effects), and participant ID was entered as a random intercept effect (Table 2, *left panel*). ActAware, 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 ActAware and Nonjudging, which uniquely explained 6.1% and 4.0% 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- $R^2 = 0.8\%$), Activity Valence (pseudo- $R^2 = 39.3\%$), and Activity Type (pseudo- $R^2 = 19.0\%$) were all significantly and positively related with state affect scores (p -values < 0.001). We further confirmed that none of the potential covariates interacted with the state-4FMQ total score in predicting state affect. Therefore, these three variables were all suitable as covariates to include in our MLM's.

In the incremental model including trait mindfulness as a covariate (Table 2, *middle panel*), the effect sizes of the state-4FMQ facets did not change from the predictive model. Though the effect of trait mindfulness was statistically significant ($p < 0.001$), it uniquely accounted for a relatively small share of variance in the model (pseudo- $R^2 = 0.6\%$).

In the robust model including Activity Valence and Activity Type as covariates (Table 2, *right panel*), the effect sizes of all state-4FMQ facets did substantially decrease from the predictive model. While much of this decrease can be attributed to the relatively strong (and unique) effect of Activity Valence (pseudo- $R^2 = 17.5\%$), the unique effect of Activity Type (pseudo- $R^2 = 2.2\%$) also played a role. Within the Activity Type factor there were notable differences based on type of activity engaged in: Resting, Social, and Physical activities were more strongly associated with state affect than Household, Cognitive, or Other activities.

Further Validation of the state-4FMQ

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-4FMQ, as was observed in the development of the state-4FMQ (Raynes & Dobkins, 2025).

Construct validity of the state-4FMQ was examined in three ways. First, we tested the multidimensionality of the state-4FMQ. We compared the model fit between the predictive model above, versus an alternative model of one total state-4FMQ 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.

Second, we tested whether the state-4FMQ was behaving in a state- rather than trait-like manner. In the incremental model above (Table 2, *middle panel*), the relationship between the state-4FMQ 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 (Table 3). As predicted, about half or more of the variation in all state-4FMQ 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-4FMQ in detecting mindful states by assessing whether the Level 2 state-4FMQ 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-4FMQ total score with a medium effect size ($F(2, 110) = 3.91, p = 0.02, \eta^2 = 0.07$). 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 = $-0.37, p = 0.08$) and past meditators (mean difference = $-0.29, p = 0.11$). While these pair-wise results did not reach statistical significance, their direction was as expected. No difference was observed between current and past meditators (mean difference = $-0.08, p = 0.92$).

Convergent validity assessed whether each state-4FMQ facet corresponded to the trait-FFMQ facet from which it derived by using bivariate correlations to assess the Level 2 state-4FMQ facets versus their corresponding trait-FFMQ facet score (obtained in Day 1). As predicted, the aligned facets of the state-4FMQ and trait FFMQ had medium to large effect sizes with the following values: Total score ($r(111) = 0.51, p < 0.001$), ActAware ($r(111) = 0.27, p = 0.003$), Observing ($r(111) = 0.41, p < 0.001$), Nonjudging ($r(111) = 0.52, p < 0.001$), and Describing ($r(111) = 0.35, p < 0.001$). Also as expected, exploratory analyses revealed that there was a medium to large relationship between Level 2 state affect versus Pre-Intervention Trait Happiness ($r(111) = 0.44, p < 0.001$).

Table 2 Variables predicting state affect during daily life experiences

Predictors	Predictive model				Incremental model				Robust model			
	<i>B</i>	<i>CI</i>	<i>p</i>	Pseudo- <i>R</i> ²	<i>B</i>	<i>CI</i>	<i>p</i>	Pseudo- <i>R</i> ²	<i>B</i>	<i>CI</i>	<i>p</i>	Pseudo- <i>R</i> ²
(Intercept)	0.58	0.49 to 0.68	< 0.001	0.061	-0.54	-1.10 to 0.02	0.059	0.061	0.37	0.26 to 0.48	< 0.001	0.018
ActAware	0.33	0.27 to 0.38	< 0.001	0.061	0.33	0.27 to 0.38	< 0.001	0.061	0.18	0.14 to 0.22	< 0.001	0.0001
Observing	0.08	0.02 to 0.14	0.009	0.003	0.08	0.02 to 0.14	0.009	0.003	-0.01	-0.06 to 0.03	0.590	0.0001
Describing	0.02	-0.05 to 0.10	0.525	0.0002	0.02	-0.05 to 0.10	0.525	0.0002	0.05	-0.01 to 0.11	0.081	0.001
Nonjudging	0.36	0.29 to 0.43	< 0.001	0.040	0.36	0.29 to 0.43	< 0.001	0.040	0.27	0.21 to 0.32	< 0.001	0.022
Trait Mindfulness					0.02	0.01 to 0.04	< 0.001	0.006	0.35	0.32 to 0.38	< 0.001	0.175
Activity Valence												0.022
Activity Type												
Household									0.07	-0.05 to 0.19	0.228	
Physical									0.36	0.22 to 0.50	< 0.001	
Other									-0.04	-0.17 to 0.10	0.597	
Restful									0.36	0.26 to 0.46	< 0.001	
Social									0.38	0.27 to 0.49	< 0.001	
σ^2	0.78				0.78				0.44			
τ_{00}	0.21	Participant			0.18	Participant			0.22	Participant		
ICC	0.21				0.18				0.34			
<i>n</i>	113	Participant			113	Participant			113	Participant		
Observations			1751				1751					
Marginal <i>R</i> ² /conditional <i>R</i> ²			0.168/0.344				0.195/0.344					0.441/0.631

Note. Left panel: model asking if the state-4FMQ facets predict Happiness. *Middle panel:* model asking if adding trait mindfulness lowers the effects of the state-4FMQ. *Right panel:* model asking if adding Activity Valence and Activity lowers the effects of the state-4FMQ. *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

Table 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-4FMQ Total	50.59	49.41	3.57
ActAware	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

^aThe 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-4FMQ was scored on a sliding scale ranging from 1 to 5. Activity Valence was scored on a Likert scale from 1 to 7

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 (Table 2) 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 of condition ($F(2, 1101) = 14.51, p < 0.001$), which was driven by the unexpected finding that the passive control group showed higher overall mindfulness than the two other groups. However, both the main effect for time ($F(2, 1101) = 2.46, p = 0.11$) and the interaction between time and condition ($F(4, 1026) = 0.80, p = 0.52$) were not significant. For trait happiness, the results of the two-way mixed ANOVA revealed a significant main effect for condition ($F(2, 1101) = 10.62, p < 0.001$). However, both the main effect for time ($F(2, 1101) = 0.44, p = 0.64$) and the interaction between time and condition ($F(4, 1101) = 0.196, p = 0.94$) were non-significant. Though not pre-registered, exploratory analyses also tested whether there were any significant differences when omitting the Active 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.

Discussion

The main finding of the current study is 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; Enkema et al., 2020; Gross et al., 2024; Killingsworth & Gilbert, 2010; 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-4FMQ), allowing us to investigate *which* aspects of state mindfulness are most closely tied to state happiness. We found the strongest unique effects of ActAware and Nonjudging on happiness, 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-4FMQ (Raynes & Dobkins, 2025), 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 weak effects of state-Observing and state-Describing on state affect (in the current study and the developmental article) may appear surprising, these findings mirror previous studies showing that trait-Observing and trait-Describing (from the trait-FFMQ) show only 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 effects 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 effects of state Observing and Describing in the current study may therefore be attributable to characteristics of our sample and should be further explored in more diverse populations. In addition, because the current study only tested state happiness, it will be fruitful for future studies to test additional theoretically guided predictive criteria, such as state stress.

Although our findings demonstrate that different facets of state mindfulness uniquely predict state affect (specifically, happiness) 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 ActAware, 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 ActAware, Nonjudging, and Observing (see Du et al., 2019 for an ESM study demonstrating a reciprocal relationship between state mindfulness and positive emotions; and Borghi et al. (2024) for a daily diary study showing several significant longitudinal bidirectional

associations, in unexpected directions, between the use and perceived helpfulness of four state mindfulness facets with daily perceived stress). Although it was not the goal of the current study to provide evidence for causality, there does exist causal evidence from previous experimental studies showing that inducing state mindfulness (through a single meditation) session increases momentary wellbeing (Bondi, 2021; Colgary et al., 2020; Liu et al., 2013). Another way to provide evidence of causality would be to conduct time-lag studies, showing that the state of one variable (either state mindfulness or happiness) at time 0 predicts the state of the other 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.

A secondary aim of the current study was to further validate the state-4FMQ. Whereas the scale development article (Raynes & Dobkins, 2025) validated the state-4FMQ in the context of a formal meditation intervention, the current study tested its validity in the context of daily, naturally occurring, mindfulness. 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 (noted above) found that state mindfulness predicts state affect, in particular happiness. Second, convergent validity of the state-4FMQ is supported by our finding of significant bivariate correlations between the Level 2 state facets and their aligned trait-FFMQ facets.

In addition to the predictive and convergent validity of the state-4FMQ observed in the current study, which corroborates the findings from the development article (Raynes & Dobkins, 2025), we also demonstrated its construct validity in three ways. First, we show that it behaves in a multidimensional fashion. Whereas the development article (Raynes & Dobkins, 2025) demonstrated multidimensionality through exploratory and confirmatory factor analysis, the current study confirmed this in the predictive model showing that incorporating the individual four facets of the state-4FMQ accounts for more variance in state happiness than is explained by the total score alone. Second, we show that the state-4FMQ behaves more like a state than a trait measure. This is evidenced by the finding that, in the predictive model, adding *trait* mindfulness as a covariate had no impact on the relationship between state mindfulness and happiness, and moreover, that trait mindfulness itself showed only a weak relationship with state happiness. In addition to corroborating an analogous validation result in the development article (Raynes & Dobkins, 2025), this finding is important 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-4FMQ is further bolstered by the results of the ICC analysis, which revealed substantial moment-to-moment variation in all state-4FMQ 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 (Medvedev et al., 2017; Truong et al., 2020).

Third, we show that the state-4FMQ demonstrates discriminant sensitivity in detecting mindful states. Even over the course of just 2 days, the state-4FMQ detected greater average (Level 2) daily mindfulness levels for current and past, than non, meditators. Larger effects may be expected among samples with more variance in the Meditation Status groupings, and if the length of administration were longer than 2 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-4FMQ), or by an experimental condition (e.g., in reference to a meditation versus a control condition, as was performed in the development article of the state-4FMQ). In both the current study and the development article of the state-4FMQ (Raynes & Dobkins, 2025), the results of this analysis by type of participant provide evidence for discriminant sensitivity.

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 2 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 produce lasting effects, even in targeted intervention studies (e.g., 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 and if the power of the resulting statistical analyses were greater.

Limitations and Future Directions

On a final note, we discuss general considerations that apply to future work employing the DRM to assess the effects of 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 experiencing during events of variable

duration from the previous day. The fact that the current, and previous, DRM studies find significant associations between variables of interest suggest that the DRM method must be reliable to some degree. While previous DRM studies have established the method's accuracy in capturing recalled *affective* states (e.g., Diener & Tay, 2014; Kahneman et al., 2004; Ludwigs et al., 2019; Schwarz et al., 2009), the present study provides novel evidence that the DRM can also reliably capture recalled *mindful* states. Furthermore, it reveals that distinct components of state mindfulness show unique associations with state affect, underscoring the DRM's sensitivity to detecting these nuanced effects. Still, because it might be difficult for participants to remember all facets of state mindfulness from the previous day, researchers might consider including a Don't Know or Not Applicable (DK/NA) response option. 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).

Related to the potential issue of unreliable (or lack of) memory, it is possible that certain events could affect the accuracy of memory in general, for example, on days where one experiences a tragedy or something spectacular (see Colombo et al., 2024), and this likely extends to any variable the participant is asked to recall on that day. One way to address this is for researchers to simply ask participants whether the day was atypical (e.g., Colombo et al., 2024; and see "daily stressors" as measured in Miller et al., 2024). In our study, we found that about 20% of participants reported that something "atypical" occurred, roughly half of which were upsetting versus wonderful. Though we had no a priori estimate, this was higher than expected and could have impacted our result. It would be ideal for future work to capture "atypicality," possibly even at the level of individual episodes, so that researchers can explore the role of this construct in greater detail.

Another issue related to the trustworthiness of DRM data stems from the general pressure the method places on participants; being asked to remember and provide detailed information about one's entire day may feel onerous. This may be especially true in studies that require participants to fill out the DRM questionnaire for multiple consecutive days, as in the current and previous (e.g., Dockray et al., 2010; Ludwigs et al., 2019) studies. In the current study, we attempted to gain some insight into whether participants felt burdened by the study procedure, with the expectation that, if they did, we would observe a decline in the number of episodes reported by participants from Day 1 to Day 2. This was not the case. Both days averaged at 15 episodes (similar to values

reported in previous DRM studies, e.g., Kahneman et al., 2004; Ludwigs et al., 2019; Schneider et al., 2020), which is consistent with the idea that completing the DRM for 2 consecutive days was not too burdensome for participants. We also found that it took participants a median of under 30 min to complete each DRM day. This is substantially less than the 45–75-min 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 be feasible to ask participants about all episodes reinstated, and therefore get a more representative dataset, without overburdening participants. We also encourage future researchers to assess participant's experience completing the DRM, such as with open-ended questions, since the DRM is less utilized than other ILD methods. Critically, asking participants about all episodes will naturally increase the size of the collected dataset and therefore increase the power of subsequent multilevel analyses. Because of the difficulty of calculating accurate simulation-based power analyses for multilevel models prior to analysis (Aguilar-Raab et al., 2021; Mathieu et al., 2012), we encourage researchers to lean towards the overcollection, rather than undercollection, of data to avoid power issues in their analyses. This can be achieved by boosting the sample size, the number of episodes collected, or both.

In addition to concerns related to the integrity of DRM data, we also raise the possibility that some of the associations observed in DRM data (or in any ILD method) might be affected by *other* aspects of one's experience during a daily episode. Covariates should be considered in any model if they affect the dependent measure, share variance with a predictor variable, or both. In the current study, trait mindfulness and Activity Type/Valence served this purpose. Future research should also consider additional covariates, such as qualities of thought during each episode, as various aspects of this construct—such as thought valence and interestingness—have been shown to influence the relationship between state mindfulness and momentary affect (Banks et al., 2016; Gross et al., 2024; Mills et al., 2021; Poerio et al., 2013; Welz et al., 2018).

The current study demonstrates that different components of state mindfulness uniquely contribute to state affect, specifically happiness, in daily life. At an applied level, these findings suggest that integrating mindfulness into daily life, particularly through Acting with 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 (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-4FMQ 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-4FMQ. Whereas the original validation of the state-4FMQ was in the context of formal mindfulness through a meditation intervention, the current study employed the state-4FMQ in the context of daily mindfulness. Like the development article, the current study demonstrates predictive, construct, and convergent validity. As such, the state-4FMQ is currently the only modification of the trait-FFMQ 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 among 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-4FMQ are promising.

Author Contribution Material preparation, data collection, and analysis were performed by SR. The design and writing of the manuscript was a joint effort between SR and KD. All authors read and approved the final manuscript.

Data Availability All data are available at the Open Science Framework (<https://osf.io/j7tfn>).

Declarations

Ethical Approval The Institutional Review Board at UCSD reviewed the study protocol prior to commencement and deemed it exempt status.

Informed Consent Informed consent was obtained from all individual participants included in the study.

Conflict of Interest The authors declare no competing interests.

Use of Artificial Intelligence Artificial intelligence was not used in the preparation of this manuscript.

Open Access This article is licensed under a Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License, which permits any non-commercial use, sharing, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if you modified the licensed material. You do not have permission under this licence to share adapted material derived from this article or parts of it. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by-nc-nd/4.0/>.

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*, *58*(12), e13937. <https://doi.org/10.1111/psyp.13937>
- 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. 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>
- 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>
- Barr, D. J., Levy, R., Scheepers, C., & Tily, H. J. (2013). Random effects structure for confirmatory hypothesis testing: Keep it maximal. *Journal of Memory and Language*, *68*(3), 255–278. <https://doi.org/10.1016/j.jml.2012.11.001>
- Bates, D., Kliegl, R., Vasishth, S., & Baayen, H. (2018). *Parsimonious mixed models*. arXiv. <http://arxiv.org/abs/1506.04967>
- 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., 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>
- 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*. Springer International Publishing. https://doi.org/10.1007/978-3-030-77644-2_24-1
- Bolger, N., & Laurenceau, J.-P. (2013). *Intensive longitudinal methods: An introduction to diary and experience sampling research*. Guilford Press.
- Bondi, T. (2021). *Mechanisms of meditation: Investigating the components and covariates of a single session of meditation* (unpublished doctoral thesis, University of California San Diego, advisor: Karen Dobkins).
- Borghi, O., Voracek, M., & Tran, U. S. (2024). Day-to-day associations between mindfulness and perceived stress: Insights from random intercept cross-lagged panel modeling. *Frontiers in Psychology*, *15*, 1272720. <https://doi.org/10.3389/fpsyg.2024.1272720>
- 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>
- Colgary, C. D., Dong, S., & Fisher, P. H. (2020). One-session mindfulness versus concentrative meditation: The effects of stress anticipation. *American Journal of Health Education*, 51(2), 120–128. <https://doi.org/10.1080/19325037.2020.1712668>
- Colombo, D., Wrzesien, M., & Baños, R. M. (2024). Recalling tough times: Exploring the effect of acceptance on autobiographical memory in daily life. *Mindfulness*, 15(3), 601–615. <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., Greenson, 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>
- Gross, M., Raynes, S., Schooler, J. W., Guo, E., & Dobkins, K. (2024). When is a wandering mind unhappy? The role of thought valence. *Emotion*. <https://doi.org/10.1037/emo0001434>
- 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), 791–802. <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>
- 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>
- 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>
- Kabat-Zinn, J. (1982). An outpatient program in behavioral medicine for chronic pain patients based on the practice of mindfulness meditation: Theoretical considerations and preliminary results. *General Hospital Psychiatry*, 4, 33–47. [https://doi.org/10.1016/0163-8343\(82\)90026-3](https://doi.org/10.1016/0163-8343(82)90026-3)
- Kabat-Zinn, J. (2003). Mindfulness-based interventions in context: Past, present, and future. *Clinical Psychology: Science and Practice*, 10(2), 144–156. <https://doi.org/10.1093/clipsy.bpg016>
- Kahneman, D., Krueger, A. B., Schkade, D. A., Schwarz, N., & Stone, A. A. (2004). A survey method for characterizing daily life experience: The day reconstruction method. *Science*, 306(5702), 1776–1780. <https://doi.org/10.1126/science.1103572>
- Kakudji, B. K., Mwila, P. K., Burger, J. R., & Du Plessis, J. M. (2020). Epidemiological, clinical and diagnostic profile of breast cancer patients treated at Potchefstroom regional hospital, South Africa, 2012–2018: An open-cohort study. *The Pan African Medical Journal*, 36, 9. <https://doi.org/10.11604/pamj.2020.36.9.21180>
- Keech, K., & Coberly-Holt, P. (2021). Strategies and tactics for multidisciplinary writing. *Journaling for Mental Health*, 2021, 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>
- Liu, X., Wang, S., Chang, S., Chen, W., & Si, M. (2013). Effect of brief mindfulness intervention on tolerance and distress of pain induced by cold-pressor task. *Stress and Health*, 29(3), 199–204. <https://doi.org/10.1002/smi.2446>
- 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, 477. <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, 2684. <https://doi.org/10.3389/fpsyg.2019.02684>
- Matuschek, H., Kliegl, R., Vasishth, S., Baayen, H., & Bates, D. (2017). Balancing type I error and power in linear mixed models. *Journal of Memory and Language*, 94, 305–315. <https://doi.org/10.1016/j.jml.2017.01.001>
- Medvedev, O., Krägeloh, C., Narayanan, A., & Siegert, R. (2017). Measuring mindfulness: Applying generalizability theory to distinguish between state and trait. *Mindfulness*, 8(4), 1036–1046. <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*, 40(5), 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, Data analysis and research publication* (Vol 3, pp. 219–241). American Psychological Association. <https://doi.org/10.1037/13621-011>
- 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>
- 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](https://doi.org/10.1108/S2046-410X(2013)0000001017)
- 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>
- Raynes, S., & Dobkins, K. (2025). *Development and initial validation of the State Four Facet Mindfulness Questionnaire*. Assessment (in press).
- 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., & 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>
- 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>
- Schwarz, N., Kahneman, D., & Xu, J. (2009). Global and episodic reports of hedonic experience. In R. Belli, F. Stafford, & D. Alwin (Eds.), *Calendar and time diary* (pp. 156–174). SAGE Publications. <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., Sanderma, 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(5), 1053–1062. <https://doi.org/10.1007/s12671-014-0354-7>
- Stone, A. A., & Shiffman, S. (1994). Ecological momentary assessment (EMA) in behavioral medicine. *Annals of Behavioral Medicine*, 16(3), 199–202. <https://doi.org/10.1093/abm/16.3.199>
- 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>
- 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>