



Development and Validation of the Mindfulness Meditation Practice Quality Scales (MMPQS) in a Clinical Population Using Rasch Analysis

Sarah Strohmaier^{1,2} · Bruno A. Cayoun³ · Alice G. Shires⁴ · Oleg N. Medvedev⁵ · Christian U. Krägeloh⁶

Received: 27 November 2025 / Accepted: 9 September 2026
© The Author(s) 2026

Abstract

Objectives With mindfulness programs and practice research increasing, measures have been developed to assess mindfulness-related constructs. However, to our knowledge, there has not yet been a measure to examine the quality of mindfulness meditation practices. This study aimed to develop and validate the Mindfulness Meditation Practice Quality Scales (MMPQS) in a clinical population.

Method The MMPQS consists of two scales: Concentration (*śamatha*) and Insight (*vipassanā*). Ten items were initially formulated for each scale. A total of 368 clinical participants undertaking Mindfulness-integrated Cognitive Behavior Therapy (MiCBT) contributed data, completing each scale immediately after the relevant practice on two measurement occasions; observations from both occasions were combined for Rasch analysis (Concentration $n = 414$; Insight $n = 313$). The Partial Credit Rasch model was applied to investigate psychometric properties of each scale as they represent two distinct dimensions.

Results Best Rasch model fit was achieved after removing misfitting items resulting in six-item versions of Concentration and Insight scales. Both showed high reliability ($PSI = 0.89–0.94$), invariance across personal factors, and unidimensionality. Construct and predictive validity of the six-item scales were supported, with particular support for divergent validity, by correlations with relevant measures in expected directions.

Conclusions This initial validation supported reliability and validity of the MMPQS as measures of Concentration and Insight practice quality, which can be utilized to support research and practice. Future research recommendations include testing the MMPQS in nonclinical populations, across more culturally heterogeneous samples, as well as other mindfulness-based programs.

Keywords Mindfulness meditation practice quality · Assessment · Rasch analysis · Psychometrics · Mindfulness-integrated cognitive behavior therapy

Mindfulness practice has been defined as the unbiased and non-reactive monitoring of phenomena spontaneously emerging in conscious awareness, with equanimity and clear understanding of their impermanent and impersonal

essence (Cayoun, 2017). Research with mindfulness-based programs (MBPs) and practices has continued to enjoy rapid growth (Baminiwatta & Solangaarachchi, 2021; Ferreira & Demarzo, 2024). Mindfulness has been found helpful for

✉ Sarah Strohmaier
sarah.strohmaier@vu.edu.au

Bruno A. Cayoun
bruno.cayoun@mindfulness.net.au

Alice G. Shires
Alice.Shires@uts.edu.au

Oleg N. Medvedev
oleg.medvedev@waikato.ac.nz

Christian U. Krägeloh
chris.krangeloh@aut.ac.nz

¹ Victoria University, Melbourne, Australia

² Canterbury Christ Church University, Canterbury, United Kingdom

³ Mindfulness-integrated CBT Institute, Hobart, Australia

⁴ University of Technology Sydney, Sydney, Australia

⁵ University of Waikato, Hamilton, New Zealand

⁶ Auckland University of Technology, Auckland, New Zealand

mental health and well-being outcomes, for both clinical (Goldberg et al., 2018) as well as non-clinical (Galante et al., 2021) populations. However, much research has focused on practice *quantity* (e.g., Strohmaier, 2020; Strohmaier et al., 2021) and less on practice *quality*, which is arguably equally important (Ribeiro et al., 2018).

To meet the needs for research to explore the increasingly diverse applications of mindfulness practice, many different measures of mindfulness have been developed, assessing a variety of mindfulness constructs, such as trait mindfulness (e.g., Baer et al., 2006), in-the-moment state mindfulness (e.g., Lau et al., 2006), and mindfulness as a skill to be learned (e.g., Walach et al., 2006). Many mindfulness measures have also been adapted for use with different age groups (e.g. Krägeloh & Strohmaier, 2024; Pallozzi et al., 2017), various clinical populations (e.g., Joos et al., 2025; Winkens, 2022), and international contexts (e.g., Henning et al., 2024). More recently, measures have begun to examine factors related to the practice of mindfulness. This includes for example formal and informal mindfulness practice adherence (Hassed et al., 2021), the assessment of a mechanism of mindfulness-based programs, such as equanimity (Rogers et al., 2021), or factors related to the participants' ability to provide a rating of following practice instructions on a visual analogue scale (Del Re et al., 2012). However, no measure so far has been developed for the quality of different types of mindfulness practices. This is important, since the quality of a practice, and not only the time spent practicing, may shape outcomes, in particular for novice meditators (Ribeiro et al., 2018; Strohmaier & Goldberg, 2024). Additionally, practice quality relating to specific types of mindfulness practices, rather than a generic measure incorporating all different practices, is crucial to ensure a more nuanced assessment.

According to Buddhist traditions, meditation includes several different practices, principally concentration (*śamatha*), insight (*vipassanā*), and loving-kindness (*metta*); whereas *śamatha* and *vipassanā* are understood as the two central practices (Khantipalo, 1995). In *śamatha* practices, the focus is on one specific object of concentration. *Vipassanā* practices involve observing and monitoring thoughts and emotions as they are to understand the true nature of reality—all mental and physical experiences are impermanent and impersonal phenomena—and *metta* practices cultivate compassion, empathy, and kindness towards oneself and others (Anālayo, 2006a). While loving-kindness is related to mindfulness practice, it serves a different purpose, focusing on a construct distinct from mindfulness, and historically is not taught in the *satipaṭṭhāna sutta*, the formal teaching on the establishment of mindfulness (Anālayo, 2006a; Sedlmeier et al., 2023).

Furthermore, Nash et al. (2013); Nash and Newberg (2023), who have developed a universal taxonomy and

classification of meditation methods, differentiate between cognitive-directed methods (which include *śamatha* and *vipassanā*) and affective-directed ones (which include *metta*). Therefore, concentration and insight practices are focused on here first and foremost, since arguably, these are the fundamental meditative practices in Buddhist traditions resulting from Pali sources (Griffiths, 1981; Hart, 1987). Although concentration and insight practices have at times been classed as “contrasting practices” (Griffiths, 1981, p. 606), these are arguably the main practices available to mindfulness meditation practitioners and thus, are practices that need to be completed with a high level of quality. While concentration and insight practices engage different processes, they are complimentary practices and become more integrated with increased meditator proficiency and experience over time (Anālayo, 2006a). The interrelationship between concentration and insight practices has been outlined in detail elsewhere (Khantipalo, 1995).

The most commonly researched mindfulness meditation practices are those pertaining to concentration and insight (Mehrmann & Rakesh, 2013). In particular, the MBPs with the most research evidence to date are Mindfulness-Based Stress Reduction (MBSR) and Mindfulness-Based Cognitive Therapy (MBCT) (Alsubaie et al., 2017), both of which employ concentration and insight (Kabat-Zinn, 1982; Segal et al., 2002). Research on other types of MBPs has also compared the effectiveness of concentration and insight meditation practices in particular, finding that different processes are involved in these practices, but that they are effective for both physiological and mental health. For example, research by Ooishi et al. (2021) found that focused attention practices (*śamatha*) elevated physiological relaxation, whereas open monitoring (*vipassanā*) practices reduced stress and elevated physiological arousal in novice meditators with a history of physical health diagnoses such as diabetes, thyroid dysfunction, or hypertension. Research on *śamatha* and *vipassanā* practices has also found these to reduce psychological distress outcomes in clinical populations, finding a decreased relapse of depressive symptoms (e.g., Kuyken et al., 2016).

While traditional psychometric approaches often assume that Likert-scale responses represent interval-level data, this assumption is frequently violated in practice, leading to measurement errors and compromised statistical analyses. The Rasch measurement model offers a robust solution to this fundamental problem by explicitly recognizing the ordinal nature of rating scale responses and providing a principled method for converting these ordinal scores into true interval measurements (Andrich et al., 2009; Siegert et al., 2010). This conversion is particularly important for mindfulness practice quality assessment, where precise measurement differences are crucial for detecting meaningful changes over time and comparing individuals across different practice contexts. The Rasch model's iterative approach

allows researchers to identify and remove misfitting items, ensuring that the final scale measures a single underlying construct with equal intervals between scale points. Additionally, the model's ability to test for differential item functioning across demographic groups ensures that the scales perform consistently across different populations, a critical consideration for measures intended for diverse clinical samples (Sutton & Medvedev, 2023). By employing Rasch analysis, researchers can develop measurement tools that meet the stringent requirements necessary for both clinical practice and research applications.

Accurately measuring the quality of practice is not only helpful for research purposes but also important for both teachers and practitioners, in particular novices, to understand both the skills to be cultivated through their meditation and the progress they are making during each practice, while at the same time highlighting any areas of difficulty where the practitioner may need further assistance.

Therefore, this study sought to develop a measure of mindfulness meditation practice quality. Due to the close relationship and the most extensive evidence for concentration and insight meditation practices, this paper aimed to develop and validate scales of these meditation practices. One of the MBPs which integrates both concentration and insight meditation practices in a measurable sequential way is Mindfulness-integrated Cognitive Behavior Therapy (MiCBT), a transdiagnostic approach principally used in clinical populations (Francis et al., 2022). Accordingly, examining the psychometric features of these scales through the implementation of MiCBT in a clinical population was both convenient and advantageous for examining their predictive validity across a wide range of moderate to severe mental health conditions. We hypothesized that mindfulness practice quality would improve with continued practice, and greater mindfulness practice quality would predict lower levels of depression, anxiety, and stress and higher satisfaction with life. We expected that mindfulness practice quality would correlate moderately with related concepts.

Method

Participants

Adults were eligible for inclusion if they were currently seeking treatment in the form of a MiCBT program (Cayoun, 2011; Cayoun & Grabovac, 2026; Cayoun et al., 2019) at either of two psychology clinic sites. In total, 368 adults met these criteria. Participants were either self-referred or referred by a medical professional to both psychology clinics. All were from a clinical population with a variety of moderate to severe mental health disorders, all diagnosed with at least one comorbid condition, as shown in Table 1.

The majority of participants were White (90.4%) Australians (91.15%), who identified as male (59.2%) with an average age of $M = 42.34$; $SD = 14.5$ (see Table 1 for further details). At the first measurement occasion, 211 participants completed the Concentration scale and 157 completed the Insight scale (Table 1). Most participants who completed the Insight scale had also completed the Concentration scale, with the smaller Insight sample reflecting attrition over the course of treatment, as insight practices are introduced later in the MiCBT program than concentration practices. Each scale was completed again on a second occasion, and observations from both occasions were combined for the Rasch analyses, yielding $n = 414$ for Concentration and $n = 313$ for Insight.

Procedure

Participation was voluntary and all participants gave informed consent prior to taking part. This study was reviewed and received ethical approval from the Canterbury Christ Church University ethics panel (ID: ETH2122-0142). They consented to their de-identified data and responses to questionnaires administered at different time points throughout treatment being utilized for the study. The data collection process is outlined in the Online Resource (Figure OR.1). Patients with psychotic or manic symptoms were excluded from the study, given the potential adverse effects that meditating can cause during these mental states (van Dam et al., 2018).

Throughout the four stages of MiCBT, several mindfulness meditation practices and cognitive-behavioral skills are taught to individuals over a standard period of 10 weeks (Cayoun, 2011; Cayoun & Grabovac, 2026; Cayoun et al., 2019), although additional time is often necessary with more severe or chronic conditions (Grabovac & Cayoun, 2025). During stage 1, meditation techniques are principally related to intrapersonal regulation. During stage 2, these are integrated with cognitive-behavioral methods, such as imaginal and in vivo exposure to reduce avoidance and promote equanimity and confidence in daily life. During stage 3, meditation techniques are integrated with cognitive-behavioral methods, such as mindful assertiveness, and applied in daily life to promote interpersonal regulation and confidence in tense interpersonal situations. During stage 4, meditation techniques are integrated with a behavioral task to help feel more connected to others and reduce the risk of relapse. Participants in this study were guided through each of these procedures.

All participants attended therapy sessions in person in direct contact with the therapist. To learn the daily meditation techniques at home, participants chose to use a set of MP3 audio tracks or identical instruction tracks embedded within *The MiCBT Guide* smartphone application to guide each practice method. Once the procedure of a method was

Table 1 Participant demographics and measure descriptive statistics for Concentration ($N=211$) and Insight ($N=157$) Scales at Time 1

Demographics/measures	Concentration				Insight			
Age: M (SD); %	41.5 (14.6)				43.17 (14.49)			
	18–35: 37.4%				18–35: 33.3%			
	36–48: 31.8%				36–48: 34%			
	49–78: 30.8%				49–78: 32.7%			
Gender (%)	Female: 40.8				Female: 39.7%			
	Male: 58.8				Male: 59.6%			
	Other: 0.5%				Other: 0.6%			
Nationality (%)	Australian: 91.9%				Australian: 90.4%			
	UK: 1.4%				UK: 1.9%			
	Other: 6.6%				Other: 7.7%			
Ethnicity (%)	White: 89.1%				White: 91.7%			
	Asian: 5.2%				Asian: 4.5%			
	Other: 5.2%				Other: 3.8%			
Measures (N)	Primary Diagnosis	<i>n</i>	Total Comorbidity	<i>n</i>	Primary Diagnosis	<i>n</i>	Total Comorbidity	<i>n</i>
	PTSD	34	4	24	PTSD	18	4	12
	OCD	18	4	17	OCD	13	4	12
	Depression	31	4	27	Depression	22	3	20
	Anxiety	50	4	30	Anxiety	37	4	26
	Personality disorder	9	2	5	Personality disorder	7	2	4
	Adjustment disorder	38	4	25	Adjustment disorder	33	4	22
	Autism	6	3	3	Autism	7	3	4
	ADHD	16	4	15	ADHD	14	4	12
	Alcohol/substance abuse	2	3	2	Alcohol/substance abuse	1	2	1
	Other	6	3	4	Other	3	2	1
Equanimity: M (SD)	45.96 (11.4)				46.94 (3.8)			
Depression: M (SD)	7.98 (4.93)				7.82 (5.08)			
Anxiety: M (SD)	6.3 (3.94)				5.86 (3.8)			
Stress: M (SD)	10.76 (3.97)				10.42 (3.93)			
Awareness: M (SD)	36.33 (5.93)				36.16 (6.21)			
Kindness: M (SD)	4.16 (0.83)				8.34 (1.68)			
Life satisfaction: M (SD)	19.61 (7.24)				20.01 (7.29)			
SEMMP: M (SD)	26.48 (6.1)				26.24 (6.13)			

M mean, *SD* standard deviation, *N* number, *PTSD* post-traumatic stress disorder, *OCD* obsessive-compulsive disorder, *ADHD* attention-deficit/hyperactivity disorder, *SEMMP* self-efficacy for mindfulness meditation practice

learned, participants practiced it in silence with a timer if they wished. See Figure OR.2 in Online Resource for details on practice data collection.

The MiCBT program integrates meditation skills in a structured and scaffolding manner, starting with a 14-min progressive muscle relaxation twice daily to motivate daily self-care and reduce stress. During that week, they are also asked to practice mindfulness of the body, remaining aware of postures and movements as often and in as many contexts as possible, as a start to keeping attention in the present moment with an easy object of focus. These two methods are generally achievable in clinical practice, unless patients are severely dissociated.

Thereafter, all mindfulness methods are practiced in 30-min sessions twice daily, starting with concentration training. For this, mindfulness of breath is introduced to teach three essential executive functions: (1) sustaining attention at the entrance

of the nostrils while holding in working memory that thoughts are just impermanent and impersonal cognitive events, (2) inhibiting the learned response (to engage with thoughts), and (3) shifting attention back to the breath without attachment to the thought. This skillset is concomitantly applied to regulate attention by cultivating mindfulness of thoughts and mental states in daily activities and shifting away from ruminative and other unhelpful thoughts when necessary.

Once sufficient ability to concentrate on the breath is acquired during meditation, a set of body scanning techniques—taught in the Burmese vipassana tradition of Ledi Sayadaw, U Ba Khin, and S. N. Goenka—are successively introduced to develop interoceptive awareness and equanimity. The primary aim in the clinical context is to improve emotion regulation through interoceptive desensitization and insight into the impermanence and impersonality of affective valence.

Finally, loving-kindness meditation is taught as part of cultivating compassion toward self and others. This is synchronously applied with explicit ethics in daily life, as a behavioral experiment, to cultivate a transpersonal awareness and sense of connectedness with others and the environment at large. When motivated by a compassionate intention, preventing harm to themselves and others helps patients prevent relapsing in a psychological disorder.

Measures

Mindfulness Meditation Practice Quality Scales (MMPQS) Development

The MMPQS consists of two scales and is designed to be administered immediately after different types of meditation practice, to measure the quality of the practice. The scales relate to Concentration and Insight mindfulness practices. The scales were developed by three members of the research team with appropriate and sufficient personal practice and teaching experience of mindfulness methods. Item development for these scales was also informed by the traditional literature on mindfulness of breath (e.g., Anālayo, 2006b; Hart, 1987) and mindfulness of body sensation taught in vipassana meditation (e.g., Anālayo, 2011, 2020, 2021; Goenka, 1998, 2000; Walshe, 2012), as applied in MiCBT (Cayoun, 2017; Cayoun, et al., 2019). Initially, a pool of items was developed and reduced to 10 items for each scale, based on primary consensus among the authors. Items on each scale were then independently rated on a 1 to 10 scale of endorsement by the authors and discussed for further screening and elaboration. The items were then improved based on these ratings and comments with any disparities discussed until a 90% consensus was reached for each item.

The Concentration scale included items such as, “I struggled to maintain my concentration”, and “I was able to refocus immediately after being distracted by mental experiences, such as thoughts and images”. The Insight scale included items such as, “When paying attention to the body, I could feel a range of pleasant, unpleasant, or neutral sensations”, and “I was able to prevent reacting to physical, emotional, or mental experiences”.

Participants were asked to indicate how truly representative each of the items on each scale is for them after having completed a practice on a 7-point Likert Scale ranging from “not at all true” (1) to “completely true” (7). Each scale was completed twice to observe progress and examine predictive validity (see Figure OR.1 for process of data collection). Concentration and Insight meditation practice quality was determined by summing all items for each scale after reverse coding. Both the Concentration ($\alpha=0.85$; $\omega=0.78$) and Insight ($\alpha=0.86$; $\omega=0.91$) scales showed good internal consistency in the current sample. The scales were designed

as stand-alone assessment tools, and which of the scales to use depends on the specific practice completed.

Equanimity Scale 16 (ES-16)

The ES-16 (Rogers et al., 2021) is a 16-item scale assessing the level of equanimity with two underlying factors, experiential acceptance and non-reactivity, on a 5-point Likert scale, with greater scores relating to greater equanimity. The ES-16 showed good internal consistency ($\alpha=0.91$; $\omega=0.90$) and has good convergent, divergent and predictive validity (Shires et al., 2023).

Depression Anxiety Stress Scale (DASS-21)

The DASS-21 (Henry & Crawford, 2005; Lovibond & Lovibond, 1995) is a 21-item scale examining general psychopathology over the past week on a 4-point Likert scale with three subscales: depression, anxiety, and stress with greater scores indicating greater symptomatology. The DASS-21 has been found to have good discriminant, convergent and construct validity, and it showed good reliability for each of subscale: depression ($\alpha=0.96$; $\omega=0.93$), anxiety ($\alpha=0.92$; $\omega=0.88$), and stress ($\alpha=0.95$; $\omega=0.89$) (Ronk et al., 2013).

Philadelphia Mindfulness Scale – Awareness (PHLMS-A)

The PHLMS-A (Cardaciotto et al., 2008) consists of ten items, which are scored on a 5-point Likert scale where greater scores demonstrate greater awareness. The awareness subscale of the PHLMS has shown good convergent and discriminant validity, and in this sample, showed good reliability ($\alpha=0.82$; $\omega=0.83$).

Compassion Scale (CS) – Kindness subscale

The kindness subscale of the CS (Pommier et al., 2019) has four items, which are scored on a 5-point Likert scale. Averaged scores are calculated for the kindness subscale, with greater scores indicating greater kindness. The CS and its subscales have shown good discriminant, construct and convergent validity and is reliable in the current sample ($\alpha=0.87$; $\omega=0.89$).

Satisfaction with Life Scale (SWLS)

The five-item SWLS (Diener et al., 1985) measures global cognitive judgements of life satisfaction on a 7-point Likert scale ranging from strongly disagree to strongly agree, with greater scores indicating greater life satisfaction. The SWLS has shown good convergent and discriminant validity and shows internal consistency ($\alpha=0.87$; $\omega=0.93$) (Jovanović et al., 2020).

Self-Efficacy for Mindfulness Meditation Practice (SEMMP-9)

The nine-item SEMMP (Birdee et al., 2020) is scored on a nine-point Likert scale with greater scores demonstrating greater self-efficacy for mindfulness meditation practice. The SEMMP is reliable ($\alpha = 0.84$; $\omega = 0.84$) in the current sample and has shown construct and convergent validity with related measures.

Data Analyses

Descriptive statistics were computed using IBM SPSS v.29 to understand characteristics of the current sample. Rasch analysis was used to validate two scales of the Meditation Practice Quality Scales, pertaining to the most common meditation practices, namely concentration and insight. Rasch analysis was conducted using RUMM2030 (Andrich et al., 2009) iteratively until the Rasch model expectations were met, including both the overall and individual item fit to the model, no local dependency, scale invariance across demographic factors and evidence of unidimensionality (Siegert et al., 2010). Each scale was completed on two measurement occasions (Figure OR.1). For the Rasch calibration, the two occasions were stacked onto a common metric rather than analyzed separately, placing person locations from both occasions on a single interval scale. This enabled direct comparison of practice quality across time (Figs. 1 and 2) and a common ordinal-to-interval conversion (Table 6), and allowed measurement invariance over time to be tested by including time as a factor in the differential item functioning (DIF) analysis. We acknowledge that the two observations contributed by each participant are not statistically independent; however, the absence of significant DIF by time for either scale indicates that item functioning was stable across occasions, supporting the appropriateness of the stacked calibration. If the likelihood-ratio test were to indicate the presence of significant differences between response options and thresholds across individual items, the unrestricted, partial-credit version of the Rasch model would be chosen (Masters, 1980), and Gustafsson's (1980) model fit criteria will be adopted as follows. Rasch model fit was assessed primarily through chi-square statistics, where both overall and individual item-trait interaction should be non-significant ($p > 0.05$), with Bonferroni adjustment applied by dividing 0.05 by the number of tests conducted (Gustafsson, 1980). Individual item fit residuals must fall within the acceptable range of -2.50 to $+2.50$ to demonstrate adequate fit to the model. While not strict requirements, optimal targeting is indicated when person location mean falls between -0.50 and $+0.50$, suggesting good coverage of the sample by the scale, and when item and person fit residuals approximate 0.00 ($SD = 1.00$) for excellent overall fit. The item location mean is automatically set to zero as the reference point in RUMM2030. Additionally, no significant differential

item functioning (DIF) should be evident across demographic factors such as gender, age, or other personal characteristics to ensure measurement invariance across different groups.

Smith's (2002) method of employing an independent-samples *t*-test comparison of person estimates from groups of items with the highest negative and positive loadings on the first principal component of the residuals. This was used to examine unidimensionality of each scale, which is supported if there are not more than 5% significant *t*-test comparisons.

Local dependency can occur when responses to one item influence the responses to another item resulting in spurious correlations thus escalating measurement error (Baghaei, 2010). Local dependency of items was evaluated employing the residual correlation matrix, which according to Christensen et al. (2017) is present when the magnitude of residual correlations exceeds the mean of all residual correlations by 0.20. Unidimensionality is affected by spurious correlations between locally dependent items, and testlets were created by combining locally dependent items. Testlets (also termed super-items) absorb the shared residual variance between dependent items, thereby resolving local dependency and satisfying strict unidimensionality without removing items or altering the content of the scale. The testlet structure is incorporated into the ordinal-to-interval conversion algorithm so that the scale continues to be scored simply by summing its individual items.

DIF for personal factors including age, gender, mindfulness practice experience, and time was examined for each scale item. Three age categories of roughly equal size were created (18–35, 36–48, 49–78 years of age). Gender was coded as either female, male, or non-binary. Mindfulness practice experience was divided into either yes or no. Time was coded as Time 1 and Time 2 of completing each scale. Nationality, ethnicity, diagnosis, and practice type, frequency and duration could not be divided into distinct equal overarching categories to calculate DIF. By employing DIF analysis, we can help ensure that this measure is robust across contexts and suitable to be applied for different populations (Sutton & Medvedev, 2023).

After Rasch analysis, ordinal scores for both scales were converted to interval scores by transforming logit scores. Construct and divergent validity of the MMPQS was examined by running bivariate correlations with several other related measures, namely the Equanimity Scale 16 (ES-16), the Depression, Anxiety, Stress Scale (DASS-21), the Philadelphia Mindfulness Scale – Awareness (PHLMS-A), the Compassion Scale – Kindness (CS), the Satisfaction with Life Scale (SWLS), and the Self-Efficacy for Mindfulness Meditation Practice (SEMMP), with a Pearson's $r < 0.50$ or greater generally considered sufficient for divergent validity (Fornell & Larcker, 1981). Predictive validity of the MMPQS scales was also completed using regression analyses with Concentration and Insight as predictors and depression, anxiety, stress, and life satisfaction as outcomes to test Hypotheses 3 and 4.

Results

Descriptive Statistics

Table 1 shows participant demographics and descriptive statistics of all measures for each scale. Participants had a variety of comorbid clinical diagnoses, with anxiety disorders being the most common primary diagnosis.

Rasch Analyses

Concentration Scale

Table 2 shows the summary of Rasch model fit statistics across the six iterative analyses of the Concentration scale (A1–A6; the final analysis, A6, includes two testlets), and Table 3 reports the item-level fit statistics for the initial 10-item and final six-item versions. Our procedure was, at each step, to identify items whose fit statistics fell outside the acceptable range (item fit residuals beyond -2.50 to $+2.50$, or local dependency in the residual correlation matrix) and then, among these statistically flagged items, to decide which to remove by considering whether removal would compromise coverage of the concentration construct. In the initial 10-item analysis (A1), four items had fit residuals outside the acceptable range: Items 1 (3.91), 2 (-3.79), 3 (2.80), and 7 (3.42) (Table 3). Items were then removed one at a time, with the model re-estimated after each removal. Item 7 (“I was so immersed in concentration that I had no thoughts, not even being aware that I am focusing on a particular object”) was removed first because, in addition to its misfit, it conceptually reflects insight rather than concentration, so its removal did not reduce coverage of the target construct; this reduced the scale to nine items (A2). Item 1 (“While I was remaining alert and focused, I felt that time was passing quickly”) was removed next, as it remained misfitting and is double-barreled and peripheral to the construct, reducing the scale to eight items (A3). Item 3 (“I felt motivated and engaged”) was then removed, as it remained misfitting and is vague and does not specifically index concentration quality, reducing the scale to seven items (A4), where a residual item-trait misfit still remained.

Inspection of the residual correlation matrix then indicated local dependency, with a high correlation ($r > 0.80$) between Item 6 (“When my attention drifted away, I was able to return quickly to my chosen object of concentration”) and Item 8 (“I was able to refocus immediately after being distracted by mental experiences, such as thoughts and images”). Because these two items measure the same refocusing ability, only one was required to preserve construct coverage; Item 6 was removed because the wording of Item 8 is preferable, providing concrete examples. This produced the six-item version (A5; Items 2, 4, 5, 8, 9, and 10), in which all item fit residuals were within the acceptable range (Table 3, final item fit), reliability was high, and the scale was invariant across personal factors. Notably, Item 2 (“I could stay focused on my chosen object of concentration”) had shown the second-largest fit residual in the initial analysis (-3.79) but was retained rather than removed, because it is the most construct-central item, directly indexing sustained attention to the object of concentration; its misfit was resolved by removing the off-construct and locally dependent items, and its fit residual was within the acceptable range in the final solution (-2.32). This illustrates our general principle of resolving misfit by removing conceptually compromised items rather than by discarding construct-central content.

When testing unidimensionality of the six-item Concentration scale resulting from A5, the independent t -test was above the 5% threshold ($A5 = 8.77\%$). When grouping the two highest positive (items 4 and 9) and two highest negative (items 5 and 8) loadings and running independent-samples t -test comparisons, this resulted in an acceptable t -test comparison of 2.5%, thus supporting strict unidimensionality when Items 4 and 9, and 5 and 8, are combined as testlets in A6. The two testlets were used only within the Rasch calibration to satisfy strict unidimensionality; the final recommended Concentration scale therefore comprises the same six individual items (Items 2, 4, 5, 8, 9, and 10) reported in Table 3, which are summed and converted to interval scores using Table 6, with the testlet structure embedded in the conversion algorithm.

Analysis of DIF found no significant difference of the six-item Concentration scale for age, gender, mindfulness

Table 2 Summary of Rasch model fit statistics for the initial (A1; 10 items), second (A2; 9 items), third (A3; 8 items), fourth (A4; 7 items), fifth (A5; 6 items), and final (A6; 6 items including testlets) analyses of the Concentration scale ($n = 414$)

Analyses	Item fit residual		Person fit residual		Goodness of fit		PSI ^a	Independent t -test
	Mean	SD	Mean	SD	$\chi^2(df)$	p		
A1	0.22	2.62	-0.38	1.37	161.36 (80)	<0.001	0.91	11.19
A2	0.28	2.62	-0.41	1.37	124.47 (72)	<0.001	0.91	12.39
A3	0.26	2.61	-0.43	1.33	121.14 (64)	<0.001	0.91	11.43
A4	0.3	1.87	-0.45	1.26	86.45 (56)	0.006	0.91	12.39
A5	0.32	1.73	-0.45	1.21	56.41 (48)	0.19	0.89	8.77
A6	-0.09	2.98	-0.45	1.03	28.86 (32)	0.63	0.91	2.5

^aPerson Separation Index. SD standard deviation, χ^2 chi square, df degrees of freedom

Table 3 Initial item-fit and final item-fit statistics of the 10-item and 6-item Concentration scale ($n = 414$)

Item	Item content	Location	SE	FitResid	χ^2
Initial item-fit					
1	While I was remaining alert and focused, I felt that time was passing quickly	0.34	0.05	3.91	14.49
2	I could stay focused on my chosen object of concentration	0.21	0.06	-3.79***	34.02
3	I felt motivated and engaged	-0.90	0.05	2.80	13.25
4 ^r	Thoughts were interfering with my ability to focus on my chosen object of concentration	0.66	0.06	1.70	7.50
5 ^r	I struggled to maintain my concentration	0.25	0.05	-1.06	10.23
6	When my attention drifted away, I was able to return quickly to my chosen object of concentration	-1.05	0.06	-2.23*	16.10
7	I was so immersed in concentration that I had no thoughts, not even being aware that I am focusing on a particular object	1.45	0.05	3.42***	35.97
8	I was able to refocus immediately after being distracted by mental experiences, such as thoughts and images	-0.55	0.05	-1.69	10.71
9	I was able to refocus immediately after being distracted by sensory experiences, such as noise and physical sensations	-0.69	0.05	0.25	9.58
10	I was able to know that I was focused without having to think about it	0.27	0.05	-1.14	9.50
Final item-fit					
2	I could stay focused on my chosen object of concentration	0.16	0.06	-2.32	14.08
4 ^r	Thoughts were interfering with my ability to focus on my chosen object of concentration	0.67	0.06	2.17	11.09
5 ^r	I struggled to maintain my concentration	0.24	0.06	-0.59	7.72
8	I was able to refocus immediately after being distracted by mental experiences, such as thoughts and images	-0.59	0.06	0.13	6.11
9	I was able to refocus immediately after being distracted by sensory experiences, such as noise and physical sensations	-0.77	0.06	2.25	11.23
10	I was able to know that I was focused without having to think about it	0.29	0.05	0.32	6.18

SE standard error, *FitResid* fit residuals, χ^2 chi square. ^rReverse-coded item; * $p < 0.05$; ** $p < 0.01$; *** $p < 0.001$

practice experience, or time. Table 3 shows the item fit statistics for the initial, 10-item version of the Concentration scale, and the item fit statistics after misfitting items have been removed, resulting in the final six-item version

of the Concentration scale. Figure 1 shows the person-item threshold distribution for Concentration items over time, clearly indicating an increase in scores from Time 1 to Time 2.

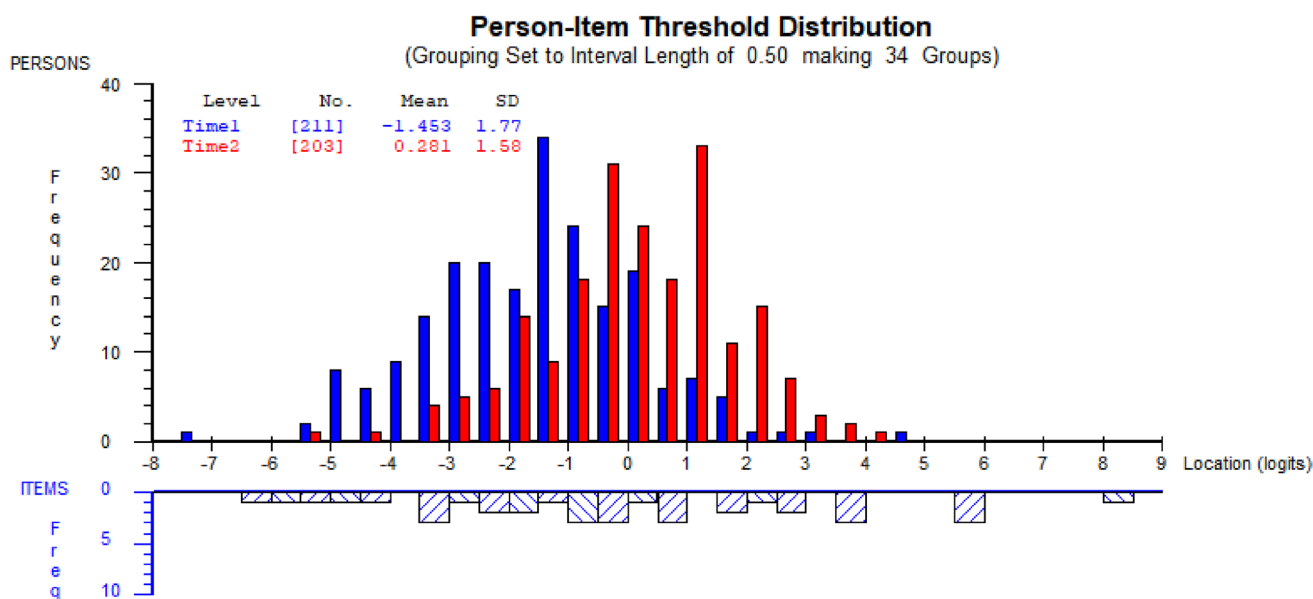


Fig. 1 Person-item threshold distribution for concentration items at Time 1 (blue) and Time 2 (red) for analysis 6 (A6)

Table 4 Summary of Rasch model fit statistics for the initial (A1; 10 items), second (A2; 9 items), third (A3; 7 items), and final (A4; 6 items) analyses of the Insight scale ($n=313$)

Analyses	Item fit residual		Person fit residual		Goodness of fit		PSI ^a	Independent t-test
	Mean	SD	Mean	SD	χ^2 (df)	<i>p</i>		%
A1	0.21	1.85	−0.63	1.63	70.92 (50)	0.03	0.94	11
A2	0.22	1.74	−0.65	1.61	43.05 (45)	0.55	0.93	12.28
A3	0.2	1.54	−0.59	1.42	32.41 (35)	0.59	0.90	5.89
A4	0.37	1.88	−0.52	1.29	29.05 (30)	0.51	0.89	3.98

^aPerson Separation Index. SD standard deviation, χ^2 chi square, *df* degrees of freedom

Insight Scale

Table 4 shows the summary of Rasch model fit statistics across the four iterative analyses of the Insight scale (A1–A4), and Table 5 reports the item-level fit statistics for the initial 10-item and final six-item versions. Unlike the Concentration scale, the initial 10-item Insight analysis (A1) already showed acceptable item fit, with all item fit residuals within the acceptable range; the subsequent removals were therefore driven by violations of unidimensionality, local dependency, and differential item functioning (DIF) rather than by item misfit. Unidimensionality was not supported in A1, with the independent *t*-test exceeding 5% when the highest positive residual loadings (Items 1, 3, and 7) were contrasted with the highest negative loadings (Items 4, 6, and 9). In addition, Item 6 (“When I did not

experience what I expected, I was able to accept it instead of letting frustration or disappointment overcome me”) showed significant DIF by age, most markedly between the youngest and oldest groups. Because the length and wording of this item may also have been unclear, and its equanimity content is captured by the remaining items, removing it did not reduce construct coverage; Item 6 was therefore removed, reducing the scale from ten to nine items (A2).

Unidimensionality was still not supported in A2 (Items 1, 2, 3, and 7 versus Items 5, 8, and 9), and the residual correlation matrix indicated local dependency involving two items. Item 1 (“I was able to feel sensations in the body”) correlated highly ($r > 0.80$) with Item 3 and shared a high residual correlation with Item 9, and Item 9 (“I was able to perceive what I experienced with objectivity and a degree of

Table 5 Initial item-fit and final item-fit statistics of the 10-item and 6-item Insight scale ($n=313$)

Item	Item content	Location	SE	FitResid	χ^2
Initial item-fit					
1	I was able to feel sensations in the body	−0.27	0.06	2.49	8.02
2	I was able to remain attentive to what I felt in the body	0.31	0.07	−1.72	8.57
3	When paying attention to the body, I could feel a range of pleasant, unpleasant, or neutral sensations	−0.06	0.06	1.03	7.78
4	I was able to prevent reacting to physical, emotional, or mental experiences	0.04	0.06	2.12	4.96
5	While paying attention to the body, I was able to accept my experiences	−0.63	0.07	−2.02	12.67
6	When I did not experience what I expected, I was able to accept it instead of letting frustration or disappointment overcome me	−0.40	0.06	−1.31	5.23
7	While paying attention to the body, I could feel the changing nature of body sensations	0.40	0.06	−0.12	5.56
8	While being attentive, I was aware that all of my experiences are impermanent	0.23	0.06	2.58	5.43
9	I was able to perceive what I experienced with objectivity and a degree of distance	0.10	0.07	−1.80	10.50
10	I could feel body sensations just as body sensations, without perceiving them as part of who I am	0.28	0.06	0.90	2.22
Final item-fit					
2	I was able to remain attentive to what I felt in the body	0.27	0.07	−1.51	7.74
3	When paying attention to the body, I could feel a range of pleasant, unpleasant, or neutral sensations	−0.07	0.06	2.45	4.70
4	I was able to prevent reacting to physical, emotional, or mental experiences	−0.01	0.06	1.20	0.82
5	While paying attention to the body, I was able to accept my experiences	−0.65	0.07	−2.24	10.94
8	While being attentive, I was aware that all of my experiences are impermanent	0.20	0.06	1.88	2.93
10	I could feel body sensations just as body sensations, without perceiving them as part of who I am	0.25	0.06	0.45	1.93

SE standard error, *FitResid* fit residuals, χ^2 chi square. * $p < 0.05$; ** $p < 0.01$; *** $p < 0.001$

distance”) correlated highly with Items 5 and 8 and is double-barreled. Because the content of both items is retained by the items with which they were redundant, Item 1 and Item 9 were removed together at this step, reducing the scale from nine to seven items (A3). This two-item removal fully accounts for the change from the nine-item A2 to the seven-item A3; no further item was removed at this step.

Unidimensionality was still not achieved in A3 (5.89%). Item 7 (“While paying attention to the body, I could feel the changing nature of body sensations”) correlated highly with Item 3 and effectively measures the same content, so it was removed, reducing the scale from seven to six items (A4). The resulting six-item version (A4; Items 2, 3, 4, 5, 8, and 10) showed acceptable item fit residuals and strict unidimensionality.

After A1, no DIF was observed for any of the person variables, such as age, gender, nationality, ethnicity, mindfulness practice experience, in A2, A3, and A4. Table 5 details the item-fit statistics for the initial, 10-item version, and the item-fit statistics of the final six-item version of the Insight scale. Figure 2 shows the person-item threshold distribution for Insight items at Time 1 and 2 for A4. The final six-item Insight scale comprehensively covers the three elements of insight, namely interoceptive awareness (Items 2 and 3), equanimity (Items 4 and 5), and awareness of impermanence and non-self (Items 8 and 10).

Conversion of Scores for MMPQS

Table 6 shows the conversion table and scores for the Concentration and Insight Scales from ordinal to interval scores for both six-item versions of the Concentration and Insight Scales.

Validity

Table 7 shows the correlation matrix for the Concentration and Insight Scales with equanimity, depression, anxiety, stress, awareness, kindness, life satisfaction, and self-efficacy for mindfulness meditation practice, demonstrating small to moderate correlations in expected directions with related measures thus indicating construct and divergent validity. The correlation between concentration and kindness is $r = 0.52$ which slightly exceeds the threshold for divergent validity of $r < 0.50$ (Fornell & Larcker, 1981). As illustrated in Table 8, Concentration practice quality significantly predicted lower levels of depression, anxiety, and stress, but did not significantly predict life satisfaction. Insight practice quality significantly predicted decreased depression, anxiety, and stress, and increased life satisfaction (Table 8). Therefore, the predictive validity of the Concentration and Insight scales is demonstrated in line with expectations.

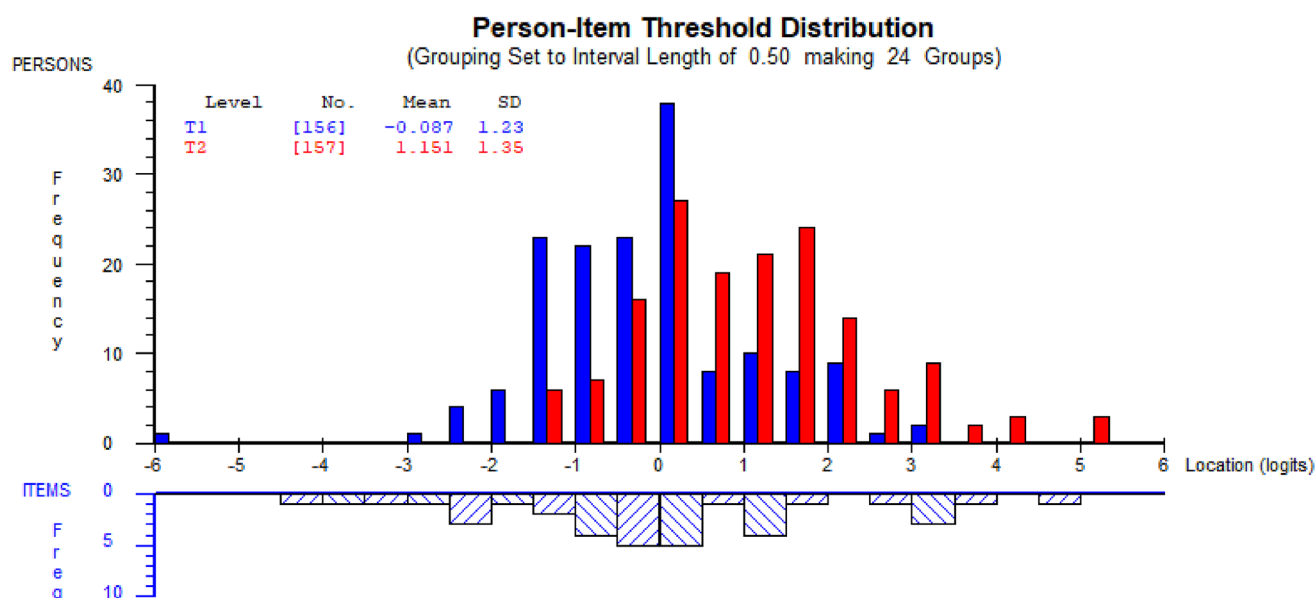


Fig. 2 Person-item threshold distribution for insight items at Time 1 (blue) and Time 2 (red) for analysis 4 (A4)

Table 6 Conversion scores for concentration and insight scales

Ordinal scores	Concentration		Insight	
	Logits	Scale	Logits	Scale
6	-7.46	6.00	-5.64	6.00
7	-6.50	8.07	-4.65	9.05
8	-5.75	9.72	-3.92	11.29
9	-5.15	11.02	-3.38	12.93
10	-4.62	12.15	-2.95	14.25
11	-4.15	13.16	-2.60	15.34
12	-3.73	14.09	-2.30	16.27
13	-3.34	14.92	-2.03	17.08
14	-2.99	15.69	-1.80	17.81
15	-2.66	16.40	-1.58	18.46
16	-2.35	17.07	-1.39	19.06
17	-2.06	17.70	-1.21	19.61
18	-1.78	18.30	-1.04	20.13
19	-1.52	18.87	-0.88	20.62
20	-1.26	19.43	-0.73	21.09
21	-1.01	19.97	-0.58	21.55
22	-0.76	20.51	-0.43	22.00
23	-0.52	21.04	-0.29	22.44
24	-0.27	21.57	-0.14	22.88
25	-0.03	22.11	0.00	23.33
26	0.23	22.65	0.15	23.79
27	0.49	23.22	0.31	24.27
28	0.76	23.81	0.47	24.77
29	1.05	24.44	0.64	25.30
30	1.36	25.10	0.83	25.87
31	1.68	25.81	1.03	26.49
32	2.03	26.57	1.25	27.17
33	2.41	27.38	1.50	27.93
34	2.82	28.27	1.78	28.79
35	3.26	29.23	2.09	29.75
36	3.74	30.27	2.44	30.82
37	4.27	31.42	2.82	31.99
38	4.87	32.70	3.25	33.30
39	5.53	34.14	3.73	34.77
40	6.32	35.85	4.30	36.52
41	7.47	38.35	5.06	38.86
42	9.16	42.00	6.08	42.00

Discussion

This study developed and validated the Mindfulness Meditation Practice Quality Scales (MMPQS) including the Concentration and Insight scales in a clinical sample of adult practitioners completing the MiCBT course. Comprehensive psychometric analyses employed the Rasch measurement model and supported six-item versions for the Concentration and Insight Scales and showed that both are reliable measures. The Concentration scale consists of

items which exclusively measure concentration with the associated cognitive flexibility (refocusing on the breath once distracted) by focusing on one specific object of concentration, having removed double-barreled and highly correlated items from the originally developed items. The six-item Concentration practice quality scale aligns well with Buddhist definitions of *śamatha* practice (Anālayo, 2006a, b). Similarly, the Insight scale assesses practice quality in terms of interoceptive awareness, equanimity, and awareness of impermanence and non-self, thus comprehensively incorporating the different aspects of *vipassanā* practice to foster understanding of the deepest nature of experience (Anālayo, 2021; Goenka, 1998; Khantipalo, 1995).

Both the Concentration and Insight scales had small to moderate correlations with measures of equanimity, depression, anxiety, stress, awareness, life satisfaction, and self-efficacy of mindfulness meditation practice, suggesting that the scales are measuring the construct they are proposing to measure, while showing sufficient divergent validity (Fornell & Larcker, 1981) in line with expectations. Findings also demonstrated good predictive validity for both scales, with concentration practice quality significantly predicting lower depression, anxiety, and stress, and insight practice quality significantly predicting decreased depression, anxiety and stress, and increased life satisfaction. Thus, these findings are in line with our expectations and previous research, suggesting that concentration and insight mindfulness meditation practices improve mental health and quality of life (Alsubaie et al., 2017; Galante et al., 2021). Furthermore, both concentration and insight practice quality improved significantly from time one to time two of completing the scales, which is consistent with our predictions and previous research suggesting improvement of mindfulness quality with continued practice (Goldberg et al., 2020).

It is worth noting that only the Insight scale showed good predictive validity for life satisfaction. Given that mindfulness of breath (concentration: *śamatha*) was taught before *vipassanā* (insight) practice, as per the MiCBT model and its corresponding lineage of teachings, it is possible that this difference is simply caused by the serial delivery of these practices—concentration practice started in week 2 and was typically practiced for one to two weeks, which was followed by insight practice, so it might be too short for patients to feel significant improvement in life satisfaction so early in therapy. While controlling for this was beyond the scope of this psychometric study, future studies could elucidate these differences by lengthening the concentration period so that both practice types are measured after the same duration.

The iterative Rasch analysis process presented several methodological challenges that were systematically

Table 7 Bivariate correlations for concentration and insight scales with related measures

Concentration Scale								
	Concentration	Equanimity	Depression	Anxiety	Stress	Awareness	Kindness	Life Satisfaction
Equanimity	0.43***							
Depression	−0.22*	−0.42***						
Anxiety	−0.27**	−0.47***	0.47***					
Stress	−0.27**	−0.65***	0.52***	0.56***				
Awareness	0.30***	0.52***	−0.16	−0.14	−0.29**			
Kindness	0.52	0.10	−0.01	0.08	−0.04	0.25**		
Life satisfaction	0.15	0.52***	−0.59***	−0.33***	−0.51***	0.30***	−0.03	
SEMMP	0.63***	−0.5***	0.33***	0.32***	0.23*	−0.46***	−0.08	−0.23*
Insight Scale								
	Insight	Equanimity	Depression	Anxiety	Stress	Awareness	Kindness	Life satisfaction
Equanimity	0.38***							
Depression	−0.37***	−0.39***						
Anxiety	−0.2*	−0.44***	0.45***					
Stress	−0.22**	−0.63***	0.47***	0.54***				
Awareness	0.39***	0.49***	−0.16	−0.06	−0.21*			
Kindness	0.06	0.08	−0.01	0.15	−0.01	0.28**		
Life satisfaction	0.26**	0.53***	−0.60***	−0.3**	−0.49***	0.30**	−0.06	
SEMMP	0.62***	−0.38***	0.22*	0.16	0.14	−0.51***	−0.17	−0.19*

SEMMP self-efficacy for mindfulness meditation practice. ^cConverted scores; * $p < 0.05$; ** $p < 0.01$; *** $p < 0.001$

Table 8 Regression results for concentration/insight practice quality predicting depression, anxiety, stress, and life satisfaction at mid-treatment

Predictor: Concentration practice quality											
Outcome	<i>R</i>	<i>R</i> ²	Adj. <i>R</i> ²	ΔR^2	<i>F</i>	<i>p</i>	<i>B</i>	<i>SE_B</i>	β	<i>t</i>	95% CI
Depression	0.22	0.05	0.04	0.05	6.22	0.01*	−0.16	0.07	−0.22	−2.49	[−0.29, −0.03]
Anxiety	0.27	0.08	0.07	0.08	9.69	0.002**	−0.14	0.04	−0.27	−3.11	[−0.22, −0.05]
Stress	0.27	0.07	0.07	0.07	9.29	0.003**	−0.17	0.06	−0.27	−3.05	[−0.28, −0.06]
Life satisfaction	0.15	0.02	0.01	0.02	2.59	0.11	0.21	0.13	0.15	1.61	[−0.05, 0.47]
Predictor: Insight practice quality											
Outcome	<i>R</i>	<i>R</i> ²	Adj. <i>R</i> ²	ΔR^2	<i>F</i>	<i>p</i>	<i>B</i>	<i>SE_B</i>	β	<i>t</i>	95% CI
Depression	0.37	0.13	0.13	0.13	17.34	<.001***	−0.24	0.06	−0.37	−4.16	[−0.35, −0.12]
Anxiety	0.20	0.04	0.03	0.04	4.9	0.03*	−0.10	0.04	−0.2	−2.21	[−0.19, −0.01]
Stress	0.22	0.05	0.04	0.05	5.98	0.02*	−0.13	0.06	−0.22	−2.44	[−0.24, −0.03]
Life satisfaction	0.26	0.07	0.06	0.07	8.08	0.005**	0.36	0.13	0.26	2.84	[0.11, 0.6]

Depression, Anxiety, Stress measured with Depression Anxiety and Stress Scale (DASS-21); Life satisfaction measured with Satisfaction With Life Scale (SWLS). *R* correlation coefficient between predictor and outcome, *R*² amount of variance accounted for by predictor, Adj. *R*² adjusted *R*²: generalizability of model, ΔR^2 adjusted *R*² change, *F* model fit of coefficient ratio of improvement, *p* significance of ΔR^2 , *B* coefficient of contribution of predictor to model showing direction and size of effect, *SE_B* Standard Error of coefficient, β standardized beta coefficient showing standard deviation change of outcome by predictor, *t* t-statistic on difference of *B* to 0, 95% C. I. 95% confidence interval

addressed through careful application of advanced psychometric techniques. Initial analyses revealed issues including disordered thresholds, local dependency between items, and violations of unidimensionality, which required strategic interventions such as item removal and creation of testlets to achieve optimal model fit. Rasch analysis allowed the ordinal scores to be converted to interval-level estimates in identifying these subtle measurement issues that would remain

undetected using traditional approaches, improving the precision of the resulting scores. These tables allow raw scores to be converted without additional software. Researchers simply locate their participant's raw summed score in the left column and read across to find the corresponding interval score. For example, a participant scoring 25 on the raw Concentration scale converts to an interval score of 22.11, while a score of 30 converts to 25.10—the equal intervals between

these converted scores now represent genuine equal units of the underlying construct, unlike the original ordinal scores where the difference between 25 and 26 may not equal the difference between 29 and 30. This conversion enables more accurate statistical analyses, precise change detection over time, and meaningful comparison of scores across different populations, ultimately enhancing both research validity and clinical decision-making.

Limitations and Future Directions

A loving-kindness subscale is planned for future inclusion in the MMPQS to enable the thorough measurement of common Buddhist meditations. However, arguably, concentration and insight practices are the most common mindfulness meditation practices in keeping with Pali sources (Anālayo, 2006a; Griffiths, 1981; Khantipalo, 1995), and according to the taxonomy by Nash et al. (2013) and Nash and Newberg (2023), include cognitive-directed meditation methods (which includes *samatha* and *vipassanā*), whereas affective-directed ones, such as loving-kindness (*metta*) meditation are classified separately. A validation of the Concentration and Insight Meditation Practice Quality Scales (MMPQS) is thus an important first step to evaluating practice quality. The validation of a loving-kindness meditation practice quality scale needs to be completed in future research.

A procedural limitation to consider is the lack of practice data for each subscale. Given the real-world clinical setting where MiCBT was delivered individually across a wide range of disorders, the frequency and duration of meditation methods were necessarily adapted by the clinician to the participant's symptom severity and attention capacity. Participants were told to continue practicing the same method for several additional days before switching to the next, as recommended in the instructional manuals (Cayoun & Grabovac, 2026; Cayoun et al., 2019). This typically resulted in some participants practicing a method more than others. Additionally, it was not possible for some participants to return to the clinic as regularly as others, which is a common occurrence in clinical settings. This also resulted in some participants practicing either of the meditation methods more than others. While this is hardly preventable, replication studies would benefit from collecting practice variability data.

Another constraint of this measure is that so far, the MMPQS has only been developed as part of MiCBT and with a clinical population of participants in a culturally homogenous sample. The fact that no significant DIF was found across demographic groups further highlights this limitation in terms of generalizability. A further limitation concerns the iterative, partly data-driven procedure used to arrive at the six-item solutions. Although each item removal was anchored to an a priori conceptual rationale as well as a statistical criterion, and the retained items preserve full

construct coverage, iterative item selection can capitalize on chance characteristics of a single sample. The item-level solutions therefore require confirmation in an independent sample before they are treated as final. Future research should therefore validate the MMPQS for use in other MBPs and with other populations, such as non-clinical, general population participants, and longer-term meditators (Kurth et al., 2023) as is the usual research strategy with newly developed measures (Krägeloh et al., 2023). Similarly, the MMPQS has been examined only with one specific length, practice amount, and face-to-face delivery of practices. Whether the MMPQS is helpful in understanding the quality of other mindfulness practice doses, i.e. practices that differ in length, frequency and amount, needs to be examined next (Strohmaier et al., 2021). Furthermore, whether there is a difference in how mindfulness is learned based on personality aspects (Strohmaier & Medvedev, 2025; Strohmaier et al., 2026) and how this relates to practice quality would be important to explore in future research.

Finally, since concentration practices were completed earlier in MiCBT than insight practices, this lag between concentration and insight as predictors on outcomes at mid-treatment needs to be noted as a methodological limitation due to its temporal confounding. Relatedly, session-to-session stability such as test–retest reliability was not assessed as a reliability criterion, because the scales are designed to detect improvement in practice quality with training and a score that remained stable across occasions would indicate insensitivity to that change. What is required for repeated use is that the items function equivalently across occasions; this was supported by the absence of differential item functioning by time for both scales, indicating that a change in score reflects a change in practice quality rather than a change in item functioning.

Overall, the MMPQS assesses concentration and insight practice quality, two central dimensions of mindfulness meditation practice, to benefit both research and practice, and provides further understanding as to whether practitioners are engaging in practices in a meaningful way and are improving in quality with greater experience.

Supplementary Information The online version contains supplementary material available at <https://doi.org/10.1007/s12671-026-02990-9>.

Acknowledgements We would like to extend our gratitude to the participants for taking part in this study.

Author Contribution SS, BAC, and CUK contributed to the study conception and design. Material preparation was performed by SS, BAC, and CUK. Data collection was completed by BAC and AGS. Data analysis was performed by SS, CUK, and ONM. The first draft of the manuscript was written by SS, and all authors commented on previous versions of the manuscript. All authors read and approved the final manuscript.

Funding Open Access funding enabled and organized by CAUL and its Member Institutions

Data Availability Materials used in this study are fully referenced. Participant permission was not sought to make raw data available. This study was reviewed and received ethical approval from the Canterbury Christ Church University ethics panel (ID: ETH2122-0142). All participants gave informed consent.

Declarations

Ethics Approval and Consent to Participate This study was reviewed and received ethical approval from the Canterbury Christ Church University ethics panel (ID: ETH2122-0142). All participants gave informed consent.

Competing interests The authors declare no competing interests.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, 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 changes were made. 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/4.0/>.

References

- Alsubaie, M., Abbott, R., Dunn, B., Dickens, C., Keil, T. F., Henley, W., & Kuyken, W. (2017). Mechanisms of action in mindfulness-based cognitive therapy (MBCT) and mindfulness-based stress reduction (MBSR) in people with physical and/or psychological conditions: A systematic review. *Clinical Psychology Review*, 55, 74–91. <https://doi.org/10.1016/j.cpr.2017.04.008>
- Anālayo, B. (2006a). *Satipaṭṭhāna: The direct path to realization* (3rd ed.). Windhorse Publications.
- Anālayo, B. (2006b). *Mindfulness of breathing: A practice guide and translations*. Windhorse Publications.
- Anālayo, B. (2011). The development of insight: A study of the U Ba Khin vipassana meditation tradition as taught by S.N. Goenka in comparison with insight teachings in the early discourses. *Fuyan Buddhist Studies*, 6, 151–174.
- Anālayo, B. (2020). Somatics of early Buddhist mindfulness and how to face anxiety. *Mindfulness*, 11, 1520–1526. <https://doi.org/10.1007/s12671-020-01382-x>
- Anālayo, B. (2021). *Deepening insight: Teachings on vedana in the early Buddhist discourses*. Pariyatti Press.
- Andrich, D., Sheridan, B., & Luo, G. (2009). *Interpreting RUMM2030*. RUMM Laboratory.
- 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, 27–45. <https://doi.org/10.1177/1073191105283504>
- Baghaei, P. (2010). A comparison of three polychotomous Rasch models for super-item analysis. *Psychological Test and Assessment Modeling*, 52(3), 343–353.
- Baminiwatta, A., & Solangaarachchi, I. (2021). Trends and developments in mindfulness research over 55 years: A bibliometric analysis of publications indexed in Web of Science. *Mindfulness*, 12, 2099–2116. <https://doi.org/10.1007/s12671-021-01681-x>
- Birdee, G. S., Wallston, K. A., Ayala, S. G., Ip, E. H., & Sohl, S. J. (2020). Development and psychometric properties of the Self-efficacy for Mindfulness Meditation Practice scale. *Journal of Health Psychology*, 25(12), 2017–2030. <https://doi.org/10.1177/1359105318783041>
- Cardaciotto, L., Herbert, J. D., Forman, E. M., Moitra, E., & Farrow, V. (2008). The assessment of present-moment awareness and acceptance: The Philadelphia Mindfulness Scale. *Assessment*, 15(2), 204–223. <https://doi.org/10.1177/1073191107311467>
- Cayoun, B. A. (2011). *Mindfulness-integrated CBT: Principles and practice*. Wiley-Blackwell.
- Cayoun, B. A. (2017). The purpose, mechanisms, and benefits of cultivating ethics in Mindfulness-integrated Cognitive Behavior Therapy. In L. Monteiro, J. Compson, & F. Musten (Eds.), *Practitioner's guide to ethics and mindfulness-based interventions. Mindfulness in behavioral health*. Springer. https://doi.org/10.1007/978-3-319-64924-5_7
- Cayoun, B. A., Francis, S. E., & Shires, A. G. (2019). *The clinical handbook of mindfulness-integrated cognitive behavior therapy: A step-by-step guide for therapists*. Wiley-Blackwell.
- Cayoun, B. A., & Grabovac, A. D. (2026). *Mindfulness-integrated CBT for well-being and personal growth: A 10-week evidence based program to cultivate inner calm, boost self-confidence and improve relationships*. Wiley-Blackwell.
- Christensen, K. B., Makransky, G., & Horton, M. (2017). Critical values for Yen's Q_3 : Identification of local dependence in the Rasch model using residual correlations. *Applied Psychological Measurement*, 41(3), 178–194. <https://doi.org/10.1177/0146621616677520>
- Del Re, A. C., Flückiger, C., Goldberg, S. B., & Hoyt, W. T. (2012). Monitoring mindfulness practice quality: An important consideration in mindfulness practice. *Psychotherapy Research*, 23(1), 54–66. <https://doi.org/10.1080/10503307.2012.729275>
- Diener, E., Emmons, R. A., Larsen, R. J., & Griffin, S. (1985). The satisfaction with life scale. *Journal of Personality Assessment*, 49(1), 71–75. https://doi.org/10.1207/s15327752jpa4901_13
- Ferreira, G. F., & Demarzo, M. (2024). Trends of research on mindfulness: A bibliometric study of an emerging field. *Trends in Psychology*, 32, 466–479. <https://doi.org/10.1007/s43076-023-00286-8>
- Fornell, C., & Larcker, D. F. (1981). Evaluating structural equation models with unobservable variables and measurement error. *Journal of Marketing Research*, 18(1), 39–50. <https://doi.org/10.2307/3151312>
- Francis, S. E. B., Shawyer, F., Cayoun, B. A., Enticott, J., & Meadows, G. N. (2022). Group mindfulness-integrated cognitive behavior therapy (MiCBT) reduces depression and anxiety and improves flourishing in a transdiagnostic primary care sample compared to treatment-as-usual: A randomized controlled trial. *Frontiers in Psychiatry*, 13, Article 815170. <https://doi.org/10.3389/fpsy.2022.815170>
- Galante, J., Friedrich, C., Dawson, A. F., Modrego-Alarcón, M., Gebbing, P., Delgado-Suárez, I., Gupta, R., Dean, L., Dalgleish, T., White, I. R., & Jones, P. B. (2021). Mindfulness-based programmes for mental health promotion in adults in nonclinical settings: A systematic review and meta-analysis of randomised controlled trials. *PLoS Medicine*, 18(1), Article e1003481. <https://doi.org/10.1371/journal.pmed.1003481>
- Goenka, S. N. (1998). *Satipaṭṭhāna sutta discourses: Talks from a course in Maha-Satipaṭṭhāna sutta*. Vipassana Research Publications.
- Goenka, S. N. (2000). *The discourse summaries: Talks from a ten-day course in Vipassana meditation (Condensed by William Hart)*. Vipassana Research Publications.
- Goldberg, S. B., Knoopel, C., Davidson, R. J., & Flook, L. (2020). Does practice quality mediate the relationship between practice time and outcome in mindfulness-based stress reduction? *Journal*

- of Counseling Psychology, 67(1), 115–122. <https://doi.org/10.1037/cou0000369>
- Goldberg, S. B., Tucker, R. P., Greene, P. A., Davidson, R. J., Wampold, B. E., Kearney, D. J., & Simpson, T. L. (2018). Mindfulness-based interventions for psychiatric disorders: A systematic review and meta-analysis. *Clinical Psychology Review*, 59, 52–60. <https://doi.org/10.1016/j.cpr.2017.10.011>
- Grabovac, A. D., & Cayoun, B. A. (2025). *The mindfulness and meditation workbook for anxiety and depression: Balance emotions, overcome intrusive thoughts, and find peace using Mindfulness-integrated CBT*. New Harbinger.
- Griffiths, P. (1981). Concentration or insight: The problematic of Theravāda Buddhist meditation-theory. *Journal of the American Academy of Religion*, 49(4), 605–614.
- Gustafsson, J.-E. (1980). Testing and obtaining fit of data to the Rasch model. *British Journal of Mathematical and Statistical Psychology*, 33(2), 205–233. <https://doi.org/10.1111/j.2044-8317.1980.tb00609.x>
- Hart, W. (1987). *The art of living: Vipassana meditation as taught by S. N. Goenka*. Harper and Row.
- Hassed, C., Flighy, A., Chambers, R., Hosemans, D., Bailey, N., Connaughton, S., Lee, S., & Kazantzis, N. (2021). Advancing the assessment of mindfulness-based meditation practice: Psychometric evaluation of the Mindfulness Adherence Questionnaire. *Cognitive Therapy and Research*, 45, 190–204. <https://doi.org/10.1007/s10608-020-10150-z>
- Henning, M. A., Joos, L., Feng, X. J., Chen, Y., Moir, F., & Webster, C. S. (2024). Southampton mindfulness questionnaire and its utility for behavioral health assessment. In C. U. Krägeloh, M. Alyami, O. N. Medvedev (Eds.), *International Handbook of Behavioral Health Assessment*. Springer. https://doi.org/10.1007/978-3-030-89738-3_19-1
- Henry, J. D., & Crawford, J. R. (2005). The short-form version of the Depression Anxiety Stress Scales (DASS-21): Construct validity and normative data in a large non-clinical sample. *British Journal of Clinical Psychology*, 44(Pt 2), 227–239. <https://doi.org/10.1348/014466505X29657>
- Joos, L., Krägeloh, C. U., Medvedev, O. N., & Henning, M. A. (2025). Acceptance and Action Questionnaire: Substance Abuse. In C. U., Krägeloh, M. Alyami, O. N. Medvedev (Eds.), *International Handbook of Behavioral Health Assessment*. Springer. https://doi.org/10.1007/978-3-030-89738-3_65-1
- Jovanović, V., Lazić, M., & Gavrilov-Jerković, V. (2020). Measuring life satisfaction among psychiatric patients: Measurement invariance and validity of the Satisfaction with Life Scale. *Clinical Psychology and Psychotherapy*, 27(3), 378–383. <https://doi.org/10.1002/cpp.2434>
- 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(1), 33–47. [https://doi.org/10.1016/0163-8343\(82\)90026-3](https://doi.org/10.1016/0163-8343(82)90026-3)
- Khantipalo, B. P. (1995). *Calm and insight: A Buddhist manual for meditators*. Routledge. <https://doi.org/10.4324/9780203565551>
- Krägeloh, C. U., Alyami, M., & Medvedev, O. N. (2023). *International handbook of behavioral health assessment (living reference work)*. Springer Nature.
- Krägeloh, C. U., & Strohmaier, S. (2024). Child and Adolescent Mindfulness Measure (CAMM) in International Contexts. In C. U. Krägeloh, M. Alyami, O. N. Medvedev (Eds.), *International Handbook of Behavioral Health Assessment*. Springer. https://doi.org/10.1007/978-3-030-89738-3_17-1
- Kurth, F., Strohmaier, S., & Luders, E. (2023). Reduced age-related gray matter loss in the orbitofrontal cortex in long-term meditators. *Brain Sciences*, 13(12), 1677. <https://doi.org/10.3390/brainsci13121677>
- Kuyken, W., Warren, F. C., Taylor, R. S., Whalley, B., Crane, C., Bondolfi, G., Hayes, R., Huijbers, M., Ma, H., Schweizer, S., Segal, Z., Speckens, A., Teasdale, J. D., Van Heeringen, K., Williams, M., Byford, S., Byng, R., & Dalgleish, T. (2016). Efficacy of mindfulness-based cognitive therapy in prevention of depressive relapse: An individual patient data meta-analysis from randomized trials. *JAMA Psychiatry*, 73, 565–574. <https://doi.org/10.1001/jamapsychiatry.2016.0076>
- Lau, M. A., Bishop, S. R., Segal, Z. V., Buis, T., Anderson, N. D., Carlson, L., Shapiro, S., Carmody, J., Abbey, S., & Devins, G. (2006). The Toronto mindfulness scale: Development and validation. *Journal of Clinical Psychology*, 62(12), 1445–1467. <https://doi.org/10.1002/jclp.20326>
- Lovibond, P. F., & Lovibond, S. H. (1995). The structure of negative emotional states: Comparison of the depression anxiety stress scales (DASS) with the Beck depression and anxiety inventories. *Behaviour Research and Therapy*, 33(3), 335–343. [https://doi.org/10.1016/0005-7967\(94\)00075-U](https://doi.org/10.1016/0005-7967(94)00075-U)
- Masters, G. N. (1980). *A Rasch model for rating scales*. University of Chicago.
- Mehrmann, C., & Rakesh, K. (2013). Principles and neurobiological correlates of concentrative, diffuse, and insight meditation. *Harvard Review of Psychiatry*, 21(4), 205–218. <https://doi.org/10.1097/HRP.0b013e31828e8ef4>
- Nash, J. D., & Newberg, A. B. (2023). An updated classification of meditation methods using principles of taxonomy and systematics. *Frontiers in Psychology*, 13, Article 1062535. <https://doi.org/10.3389/fpsyg.2022.1062535>
- Nash, J. D., Newberg, A. B., & Awasthi, B. (2013). Toward a universal definition and taxonomy for meditation. *Frontiers in Psychology*, 4, 806. <https://doi.org/10.3389/fpsyg.2013.00806>
- Ooishi, Y., Fujino, M., Inoue, V., Nomura, M., & Kitagawa, N. (2021). Differential effects of focused attention and open monitoring meditation on autonomic cardiac modulation and cortisol secretion. *Frontiers in Physiology*, 12, Article 675899. <https://doi.org/10.3389/fphys.2021.675899>
- Pallozzi, R., Wertheim, E., Paxton, S., & Ong, B. (2017). Trait mindfulness measures for use with adolescents: A systematic review. *Mindfulness*, 8, 110–125. <https://doi.org/10.1007/s12671-016-0567-z>
- Pommier, E., Neff, K. D., & Tóth-Király, I. (2019). The development and validation of the compassion scale. *Assessment*. <https://doi.org/10.1177/1073191119874108>
- Ribeiro, L., Atchley, R. M., & Oken, B. S. (2018). Adherence to practice of mindfulness in novice meditators: Practices chosen, amount of time practiced, and long-term effects following a mindfulness-based intervention. *Mindfulness*, 9(2), 401–411. <https://doi.org/10.1007/s12671-017-0781-3>
- Rogers, H. T., Shires, A. G., & Cayoun, B. A. (2021). Development and validation of the Equanimity Scale-16. *Mindfulness*, 12, 107–120. <https://doi.org/10.1007/s12671-020-01503-6>
- Ronk, F. R., Korman, J. R., Hooke, G. R., & Page, A. C. (2013). Assessing clinical significance of treatment outcomes using the DASS-21. *Psychological Assessment*, 25(4), 1103–1110. <https://doi.org/10.1037/a0033100>
- Sedlmeier, P., Beckel, A., Conrad, S., Husmann, J., Kullrich, L., Lange, R., Müller, A.-L., Neumann, A., Schaaf, T., Schaub, A., Tränkner, A., & Witzel, B. (2023). Mindfulness meditation according to the satipathana sutta: A single-case study with participants as collaborators. *Mindfulness*, 14, 1636–1649. <https://doi.org/10.1007/s12671-023-02160-1>
- Segal, Z. V., Williams, M. G., & Teasdale, J. D. (2002). *Mindfulness-based cognitive therapy for depression: A new approach to preventing relapse*. Guilford Press.
- Shires, A., Osborne, S., Cayoun, B. A., Williams, E., & Rogers, K. (2023). Predictive validity and response shift in the Equanimity Scale-16. *Mindfulness*, 14, 2880–2893. <https://doi.org/10.1007/s12671-023-02257-7>

- Siebert, R. J., McPherson, K. M., & Tennant, A. (2010). Rasch analysis of the World Health Organization Quality of Life instruments. *Value in Health*, 13(2), 275–283.
- Smith, E. V., Jr. (2002). Detecting and evaluating the impact of multidimensionality using item fit statistics and principal component analysis of residuals. *Journal of Applied Measurement*, 3(2), 205–231.
- Strohmaier, S. (2020). The relationship between doses of mindfulness-based programs and depression, anxiety, stress, and mindfulness: A dose-response meta-regression of randomized controlled trials. *Mindfulness*, 11(6), 1315–1335. <https://doi.org/10.1007/s12671-020-01319-4>
- Strohmaier, S., & Goldberg, S. B. (2024). Longitudinal increases in mindfulness practice quality are associated with changes in psychological outcomes and not vice versa – A brief report. *Current Psychology*, 43, 18517–18520. <https://doi.org/10.1007/s12144-024-05644-y>
- Strohmaier, S., Jones, F. W., & Cane, J. E. (2021). Effects of length of mindfulness practice on mindfulness, depression, anxiety, and stress: A randomized controlled experiment. *Mindfulness*, 12, 198–214. <https://doi.org/10.1007/s12671-020-01512-5>
- Strohmaier, S., Jones, F., & Medvedev, O. (2026). Evaluating the effectiveness of mindfulness practice doses in the general population through personalization-A randomized controlled experiment. *Mindfulness*, 17, 731–745. <https://doi.org/10.1007/s12671-026-02762-5>
- Strohmaier, S., & Medvedev, O. M. (2025). A latent profile analysis of the Big Five personality and mindfulness traits in the general population. *Personality and Individual Differences*, 245, Article 113287. <https://doi.org/10.1016/j.paid.2025.113287>
- Sutton, A., & Medvedev, O. N. (2023). Development and validation of the Awareness Outcomes Measure (AOM) using Rasch approach. *Mindfulness*, 14, 473–481. <https://doi.org/10.1007/s12671-022-02047-7>
- van Dam, N. T., van Vugt, M. K., Vago, D. R., Schmalzl, L., Saron, C. D., Olendzki, A., Meissner, T., Lazar, S. W., Kerr, C. E., Gorchov, J., Fox, K. C. R., Field, B. A., Britton, W. B., Brefczynski-Lewis, J. A., & Meyer, D. E. (2018). Mind the hype: A critical evaluation and prescriptive agenda for research on mindfulness and meditation. *Perspectives on Psychological Science*, 13(1), 36–61. <https://doi.org/10.1177/1745691617709589>
- Walach, H., Buchheld, N., Büttenmüller, V., Kleinknecht, N., & Schmidt, S. (2006). Measuring mindfulness – The Freiburg mindfulness inventory (FMI). *Personality and Individual Differences*, 40, 1543–1555. <https://doi.org/10.1016/j.paid.2005.11.025>
- Walshe, M. (2012). *The long discourses of the Buddha: A translation of the Digha Nikaya*. Wisdom Publications.
- Winkens, L. H. H. (2022). Mindful Eating Behavior Scale (MEBS). In O. N. Medvedev, C. U. Krägeloh, R. J. Siebert, N. N. Singh (Eds.), *Handbook of Assessment in Mindfulness Research*. Springer. https://doi.org/10.1007/978-3-030-77644-2_34-1

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.