ORIGINAL PAPER

Validation and Factor Structure of the Mindfulness-Based Self Efficacy Scale-Revised

Bruno Cayoun[1,2] · Bradley Elphinstone[3] · Natasha Kasselis[2] · Glenn Bilsborrow[1,2] · Clive Skilbeck[2]

Abstract

Objectives The initial version of the Mindfulness-based Self-Efficacy Scale (MSES) was constructed to assess changes in the perception of self-efficacy in clinical populations undertaking mindfulness training as part of therapy. Although it has been used increasingly since its subsequent revision a decade ago, the factor structure of this 22-item revised version (MSES-R) has not yet been confirmed. The current study investigated the factor structure and measurement invariance of the MSES-R in heterogenous clinical samples, and sensitivity to change over the course of a mindfulness-based therapy program.

Methods Confirmatory factor analysis was performed on the MSES-R scores collected online from two Australian samples (clinical $N = 1378$; community $N = 2866$), two Canadian samples (clinical $N = 595$; community $N = 321$), and one Australian university student sample ($N = 521$). Other questionnaires were used to further assess convergent and discriminant validity.

Results The MSES-R provided adequate fit across all samples, although factors containing fewer items had low reliability. The MSES-R displayed scalar measurement invariance between the clinical and community Australian samples, and across the Australian and Canadian samples. Both community samples score significantly higher on the MSES-R than the clinical samples. In the student sample, respondents who regularly engaged in meditative or contemplative practices scored higher on the MSES-R than non-meditators. The MSES-R was also associated with higher scores on the Five Facet Mindfulness Questionnaire and with lower levels of psychological distress in the student sample. In a separate clinical sample ($N = 68$), MSES-R scores increased significantly over the duration of a 10-week mindfulness-based therapy.

Conclusions The current findings provide additional evidence that the MSES-R is a useful addition to the battery of mindfulness-based assessment tools to investigate the efficacy and outcomes of mindfulness-based programs, including in clinical settings.

Keywords MSES · Self-efficacy · Mindfulness scale · Mindfulness questionnaire · Mindfulness measures

An increasing number of psychological interventions involving the development of mindfulness skills through daily meditation practice have emerged over the last two decades (see Zhang et al., 2021 for a review). Outcome studies of therapies requiring daily meditation show that mindfulness practice is associated with increased well-being and improved quality of life in both clinical and non-clinical populations (see Gu et al., 2015 for a review). The cultivation of mindfulness has been informed by Buddhist contemplative traditions, and its practice in modern secular applications is defined in various ways, in accordance with teachers and authors' own training lineage or clinical orientations. The practice of mindfulness is characterized by directing unbiased attention to one's experience in the present moment and the quality of this attention or awareness is critical (Block-Lerner et al., 2005).

Kabat-Zinn (2003) defined mindfulness as "the awareness that emerges through paying attention on purpose, in the present moment, and nonjudgmentally to the unfolding of experience" (p. 145). Others specify that mindfulness practice includes total openness and conscious acceptance of all thoughts, feelings, and sensations that arise, regardless of whether they are positive, negative, pleasant,

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

[1] MiCBT Institute, 277 Macquarie Street, Hobart, TAS 7000, Australia

[2] School of Psychology, University of Tasmania, Hobart, TAS, Australia

[3] School of Psychology, Swinburne University of Technology, Hawthorn, VIC, Australia

or painful (Bishop et al., 2004). Another description suggests that equanimity is necessary to prevent reactivity and biased attention, and for mindfulness meditation to be differentiated from mere attention training (Cayoun et al., 2019; Eberth et al., 2019; Hart, 1987). Effort is being made to systematically measure equanimity (e.g., Rogers et al., 2020). Equanimity has been defined as "a balanced reaction to joy and misery, which protects one from emotional agitation" (Bodhi, 2005, p. 154), and in secular scientific contexts as an "even minded mental state or dispositional tendency towards all experiences or objects, regardless of their affective valence (pleasant, unpleasant or neutral) or source" (Desbordes et al., 2015, p. 357).

Accordingly, cultivating equanimity promotes one's greater ability to regulate emotion and tolerate distress. In turn, greater coping ability resulting from increased equanimity improves one's sense of self-efficacy in facing common stressors. In Buddhist psychology, the purpose of cultivating mindfulness and equanimity is to progressively develop an ever-deeper level of insight into the fundamental mechanisms of the human experience (Bodhi, 2005). Importantly, the "stages of insight" (Buddhaghosa, 2010), which a meditator progressively acquires through mindfulness practice, are typically measured through their integration of meditative insight into daily life, including trust in one's ability to self-regulate across various domains of experience (Bergomi et al., 2012). For instance, as a person's insight into the impermanence of all phenomena develops, decreased attachment to thoughts (e.g., expectations), body sensations, and objects progressively leads to facing life's vicissitudes with greater ease (Hart, 1987). This includes, but is not limited to, greater ease with traversing difficult emotions, tolerating distress, interacting socially and intimately, and ease with taking responsibility for one's actions. Nonetheless, this integration of skills derived from the progressive effects of mindfulness is seldom captured by questionnaires which attempt to directly assess individual differences in trait mindfulness.

It has also been pointed out that many self-report questionnaires attempting to measure mindfulness may not align with the original Buddhist teachings of mindfulness (Feng et al., 2018) and are limited in their ability to measure the effectiveness of mindfulness practice in facilitating a meditator's insight (Van Dam et al., 2018). Across several measures, respondents may interpret and respond differently to items based on their different levels and understanding of mindfulness. To respond accurately to questions about being mindful requires adequate mindfulness, as it is inherently difficult to self-report about states of which one is unaware. This means that as people develop insight through mindfulness practice, they are more able to detect when they are not mindful, thereby moderating or even minimizing their scores on

mindfulness (Grossman & Van Dam, 2011). An alternative approach to measuring the effectiveness of mindfulness following an intervention is to assess its consequences on one's perceived self-efficacy in overcoming daily stressors. One of the self-report questionnaires used to measure the resulting benefits of mindfulness training is the Mindfulness-based Self-Efficacy Scale-Revised (MSES-R).

The view that the principal purpose of mindfulness practice is to reduce suffering, rather than just "being mindful," is well documented (e.g., Kabat-Zinn, 2006; Segal et al., 2002). Increased confidence in one's ability to overcome barriers to one's insight and well-being is also a well-documented measure of meditative progress in the stages of insight taught in Buddhist teachings (Buddhaghosa, 1990; see also Grabovac, 2015). In line with this understanding, the original MSES was developed with a focus on measuring skills that people felt improved in their lives as a consequence of mindfulness practice. It was developed to measure *the confidence in achieving the original purpose of mindfulness*, rather than measuring its construct. This is congruent with studies that have highlighted links between mindfulness and several forms of self-efficacy for improving self-regulation. For example, Luberto et al. (2013) demonstrated that coping self-efficacy partially explains the relationship between mindfulness and emotion regulation difficulties. Thus, the revised measure, the MSES-R, is expected to be particularly helpful in a clinical context because a strong sense of self-efficacy (i.e., a person's perception or belief in their ability to perform certain skills or act effectively to attain their goals; Bandura, 1997) is related to greater effort, persistence, and self-benefitting behaviors (Schwarzer, 2008).

There are two main ways of conceptualizing self-efficacy in the context of mindfulness. One is the person's perception or belief in their ability to practice mindfulness skillfully. For example, Chang et al. (2004) developed a self-report measure that assesses how self-efficacious a person feels in practicing mindfulness skills. The other way of conceptualizing self-efficacy, which applies to the MSES-R, in the context of a mindfulness-based intervention (MBI) is the person's perception or belief in their ability to reduce suffering and improve their daily experiences *as a result* of mindfulness practice. Because it was specifically designed as an outcome measure, the MSES-R assesses skills which typically develop from becoming more mindful while confronted by common stressors in daily life. Accordingly, improved self-efficacy in the six areas of functioning assessed by the MSES-R (i.e., Emotion Regulation, Equanimity, Social Skills, Distress Tolerance, Taking Responsibility, and Interpersonal Effectiveness), especially in relation to their interoceptive features, is commonly observed and likely to be generalizable following MBIs which emphasize insight (Cayoun et al., 2019; Nyklíček, 2020).

While no evidence of factorial validity has been published for the 35-item MSES, the factorial validity of the 22-item MSES-R has only been displayed in an unpublished study of an Australian student sample (Kasselis, 2011) and a Turkish translation of the measure administered to a non-clinical student sample (Atalay et al., 2017). In Kasselis' EFA study, the 22 retained items had a loading above 0.5, accounting for 57.5% of the total variance. The total and each subscale of the MSES-R distinguished participants with a current mental illness from those without (MSES-R total $p < 0.01$, $g = 0.79$). All subscales of the MSES-R also showed significant positive correlations with the FFMQ (see Kasselis, 2011, for more detail). Atalay et al.'s study included 713 students in 5th, 6th, and 7th grades from two different public schools. Internal consistency was acceptable (Cronbach's $\alpha = 0.72$) and the factors Emotion Regulation (0.73), Equanimity (0.68), Social Skills (0.65), Distress Tolerance (0.62), Taking Responsibility (0.61), and Interpersonal Effectiveness (0.65) were found to be acceptable after the low number of items in each subscale was taken into consideration.

Studies investigating improvements following MBIs provided support for the validity of both versions of the MSES. For instance, using the 35-item measure, Alexander et al. (2012) assessed the effectiveness of cognitive or mindfulness training programs in individuals with diagnoses of mood disorders. Participation in either form of training was associated with a significant increase in MSES scores and reductions in depressive symptoms over a 3-month period. In a sample of individuals with bipolar disorder, Van Dijk et al. (2013) found that scores on the 35-item MSES increased through participation in a 12-week Dialectical Behavior Therapy program, which included aspects of mindfulness training. While scores significantly increased for respondents in the intervention and control groups, the increase was significantly more pronounced in the intervention group.

Two published studies have used the 22-item measure. Goldstein et al. (2018) allocated individuals to an 8-week MBSR program, matched exercise condition, or a wait list. A structural equation model indicated that the MSES-R and mindfulness as measured by the Mindful Awareness Attention Scale (MAAS; Brown & Ryan, 2003) were positively correlated. Higher scores on both measures were directly associated with reduced self-reported stress and increased general mental well-being. Importantly, the MSES-R, but not the MAAS, appeared to uniquely mediate the relationship between participation in the MBSR program with both reduced stress and greater general mental health. Voith et al. (2020) found higher scores on the 22-item MSES-R to be associated with reduced symptoms of post-traumatic stress disorder, adverse childhood disorders, and symptoms associated with experiencing trauma. Hence, both versions of the MSES demonstrated sensitivity to change in clinical samples.

Therefore, rather than focusing on theory and practice of mindfulness, this study examined missing psychometric evidence of the MSES-R to better inform users in the field. Although previous studies have suggested good criterion validity, the factorial validity of the six-factor MSES-R in clinical samples and Western community samples remains unexplored. Moreover, it is not clear whether respondents in clinical and non-clinical populations respond similarly to MSES-R questions.

The aim of the first study was therefore to investigate the model fit of the MSES-R, and the measurement invariance across community and clinical samples. It was also hypothesized that clinical samples would report lower scores on the MSES-R than non-clinical samples. To further examine the validity of the MSES-R, the second study examined the relationship between the MSES-R and the Five Facet Mindfulness Questionnaire (FFMQ). It was expected that the MSES-R would be positively correlated with the FFMQ. We also investigated potential differences between meditators and non-meditators, and expected that meditators would score higher than non-meditators on the FFMQ and MSES-R. In the third study, it was expected that MSES-R scores in a clinical sample would increase as a result of participating in an MBI. In both studies 2 and 3, it was expected that higher MSES-R scores would be associated with lower levels of symptoms of depression, anxiety, and stress.

Study 1

Method

Participants

Four samples, totaling 5160 participants, were used to assess the model fit of the MSES-R. These were an Australian clinical ($n = 1378$; age range: 18–84, $M = 40.51$, $SD = 13.18$; male $n = 431$, female $n = 947$), Australian community ($n = 2866$; age range: 18–91, $M = 41.90$, $SD = 13.21$; male $n = 823$, female $n = 2043$), Canadian clinical ($n = 595$; age range: 18–78, $M = 40.35$, $SD = 13.32$; male $n = 156$, female $n = 439$), and a Canadian community ($n = 321$; age range: 18–87, $M = 39.85$, $SD = 13.48$; male $n = 72$, female $n = 249$) sample. Census data collected in 2018, which was within a similar timeframe to the collection of the data in the current study, indicated that the average age in Australia was 38.88 (median = 37.31; Australian Bureau of Statistics, 2021). The median age in Canada during the same approximate period was 40.8 years (Statistics Canada, 2019). In both countries, the population was approximately 51% female. Thus, the samples in the current study were roughly representative for age but not gender.

Procedure

All respondents freely accessed a website hosting the 22-item MSES-R. These respondents provided basic demographic information (i.e., age, gender) and information about their previous and present mental health, including whether they had received a clinical diagnosis of any mental disorder or illness, and the presence of current psychological symptoms. This information was used to classify people into clinical (i.e., current mental illness or disorder) or community (i.e., no current diagnosis, treatment, or clinical symptoms) samples. However, specific diagnostic information was not collected.

To initially develop the MSES, the data pool to develop items was collected from several hundred psychiatric patients in private and public hospitals, and community health agencies between 2001 and 2004. During this time, all patients participated in an 8-week mindfulness meditation program. Based on the participants' description of their day-to-day experiences during the weekly mindfulness training sessions and at various follow-ups, 1300 statements were collected and organized into seven categories:

(i.e., behavioral, cognitive, interoceptive, affective, interpersonal, and two general facets, avoidance, and mindfulness). The five most recurrent types of statements within each category were retained and formed an early 35-item scale (Cayoun, 2011). Items were constructed using common language that could be understood by all, not just those who had prior experience of mindfulness. A 22-item version MSES-Revised (MSES-R) was developed (Kasselis, 2011), with shortened subscales and the removal of two facets. The remaining facets were Emotion Regulation, Equanimity, Social Skills, Distress Tolerance, Taking Responsibility, and Interpersonal Effectiveness.

Measures

MSES-R Items (e.g., "I get easily overwhelmed by my emotions") are measured on a 5-point Likert scale from 0 (not at all) to 4 (completely). Respondents are provided with the prompt, "Select a response according to how much you agree with each statement below. Try not to spend too much time on any one item. There are no right or wrong answers." All items are shown in Table 1.

Table 1 The 22-item MSES-R

Emotion Regulation

1: I get easily overwhelmed by my emotions

2: When I feel very emotional, it takes a long time for it to pass

3: It is often too late when I realize I overreacted in a stressful situation

4: I get so caught up in my thoughts that I end up feeling very sad or anxious

5: I get caught up in unpleasant memories or anxious thoughts about the future

6: When I have a problem, I tend to believe it will ruin my whole life

Social Skills

7: I find it difficult to make new friends

8: I try to avoid uncomfortable situations even when they are really important

9: I can feel comfortable around people

Equanimity

10: I feel comfortable saying sorry when I am in the wrong

11: I can face my thoughts, even if they are unpleasant

12: I can deal with physical discomfort

13: When I feel physical discomfort, I relax because I know it will pass

Distress Tolerance

14: When I have unpleasant feelings in my body, I prefer to push them away

15: I avoid feeling my body when there is pain or other discomfort

16: I do things that make me feel good straightaway even if I will feel bad later

Taking Responsibility

17: My actions are often controlled by other people or circumstances

18: Seeing or hearing someone with strong emotions is unbearable to me

19: If I get angry or anxious, it is generally because of others

Interpersonal Effectiveness

20: I can resolve problems easily with my partner (or best friend if single)

21: I feel I cannot love anyone

22: I am often in conflict with one (or more) family member

Data Analyses

Model Testing Analyses were conducted with JASP (Version 0.14.1; JASP Team, 2020). The Confirmatory Factor Analyses (CFA) used the DWLS estimator and robust estimation of errors. This is based on recommendations provided by Sellbom and Tellegen (2019; see also Li, 2016), who argue that standard Likert-type response formats may be best considered ordinal rather than continuous measures. While Sellbom and Tellegen suggest that up to four response categories (e.g., 1 = strongly disagree, 4 = strongly agree) may be best analyzed as ordinal data, they also recommend that robust maximum likelihood estimation be used cautiously with more categories and if the data approaches normality. The authors highlight that the inappropriate use of maximum likelihood estimation may result in 21 appearing to be lower than should be considered the case, leading to needlessly rejecting or modifying an otherwise appropriate model. Given that five response options, as is the case for the MSES-R, can be arguably considered as appropriate for ordinal analyses (Shi & Maydeu-Olivares, 2020) and the data is not expected to be normally distributed, DWLS appeared to be a suitable estimator available in JASP. Nonnormality was expected given the potential for some items to generally receive responses above or below the mid-point on the 5-point response scale. Accordingly, across the 22 MSES-R items, because of z-scores exceeding a threshold of $z \pm 3.29$ ($p < 0.001$), 12 items were significantly skewed.

In order to determine model fit, the standard criteria (i.e., CFI and TLI ≥ 0.95, RMSEA ≤ 0.06, and SRMR ≤ 0.08; Hu & Bentler, 1999) were considered. It must be noted that DWLS tends to result in inflated CFI and TLI, and lower RMSEA values than would be the case if maximum likelihood estimation was used (Xia & Yang, 2019). According to Shi and Maydeu-Olivares (2020), SRMR is robust to estimation methods, and was therefore given primary consideration.

Invariance Testing To determine the appropriateness of comparing groups on MSES-R scores, the measure was assessed for measurement invariance across the clinical and community subsamples within and across the Australian and Canadian samples. As described by Van de Schoot et al. (2012), measurement invariance involves assessing nested models with increasingly strict constraints. This process commences with configural invariance in which a model with no constraints is examined to ensure that the model provides adequate fit across groups. Next, metric invariance examines whether respondents interpreted items in a similar manner by holding factor loadings constant across both groups. Scalar invariance subsequently holds all factor loadings and intercepts to be equal, and indicates whether both groups scored similarly on the items. This level of invariance suggests that it is appropriate to compare mean scores across groups.

While differences in χ^2 can be used to determine significant differences between models, χ^2 is sensitive to larger sample sizes (Cheung & Rensvold, 2002). Therefore, we prioritized the recommendations provided by Chen (2007), by which evidence of invariance is observed if $\Delta \text{CFI} \leq -0.01$, and $\Delta \text{SRMR} \leq 0.01$ or $\Delta \text{RMSEA} \leq 0.015$. The differences in χ^2 are presented in Table 2, along with changes in CFI, SRMR, and RMSEA.

Comparisons Between Groups A 2×2 ANOVA was considered to be an appropriate method of investigating the potential main effects of country (i.e., differences between the Australian and Canadian samples), and sample type (i.e., community and clinical samples), as well as any potential interaction. Separate 2×2 ANOVAs were also considered to examine differences on each of the six MSES-R facets, pending an assessment of the reliability of each facet to determine whether such analyses would be worthwhile. Reliability was examined using McDonald's ω, which is considered to be more accurate than Cronbach's α (McNeish, 2018).

Table 2 Model fit results for Australian and Canadian clinical and community samples, and an Australian student sample

	$\chi^2(203)$	CFI	TLI	SRMR	RMSEA (90% confidence interval)
Australia clinical ($N = 1379$)	804.510***	.962	.957	.049	.046 (.043–.050)
Australia community ($N = 2867$)	1231.741***	.979	.976	.043	.042 (.040–.044)
Canada clinical ($N = 595$)	354.111***	.976	.973	.049	.035 (.029–.041)
Canada community ($N = 321$)	206.246	.999	.999	.051	.007 (.000–.025)

*$p < .05$; *** $p < .001$

Results

Confirmatory Factor Analysis

Table 2 displays the fit indices in the Australian and Canadian clinical and community subsamples. As the construct of mindfulness-based self-efficacy is considered to comprise the six facets previously identified (Atalay et al., 2017; Kasselis, 2011), the tested model comprised the six-factor loading on an overall factor representing mindfulness-based self-efficacy. Model fit appeared to be acceptable in each analysis. Even with the consideration of DWLS estimation inflating CFI and TLI and decreasing RMSEA, the SRMR results were indicative of good model fit in all analyses. The model in the Australian sample is shown in Fig. 1, and the Canadian sample model in Fig. 2.
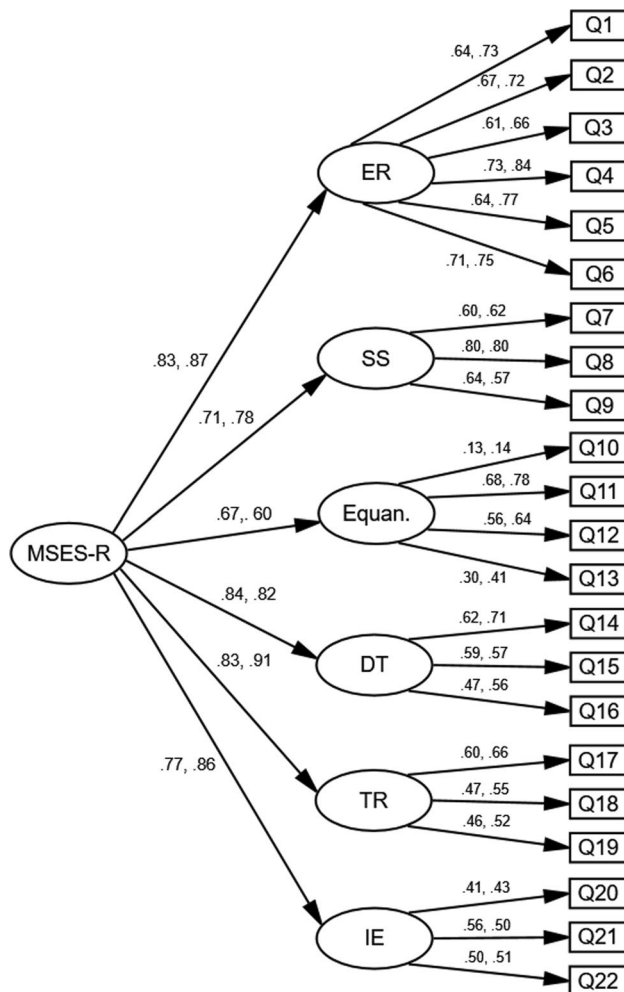
Measurement Invariance

Within each of the Australian and Canadian samples, the clinical and community samples were analyzed for invariance, as well as cross-country invariance by assessing the invariance of both clinical and also both community samples. As shown in Table 3, based on significant differences in χ^2 between the configural and metric models, the models were not invariant. However, based on Chen's (2007) criteria, the differences in CFI, SRMR, and RMSEA were indicative of metric and scalar invariance, with one exception. Specifically, the difference in CFI for the metric model in the clinical and community Canadian samples narrowly exceeded the threshold; however, the results for SRMR and RMSEA were in line with the evidence needed to support metric invariance.
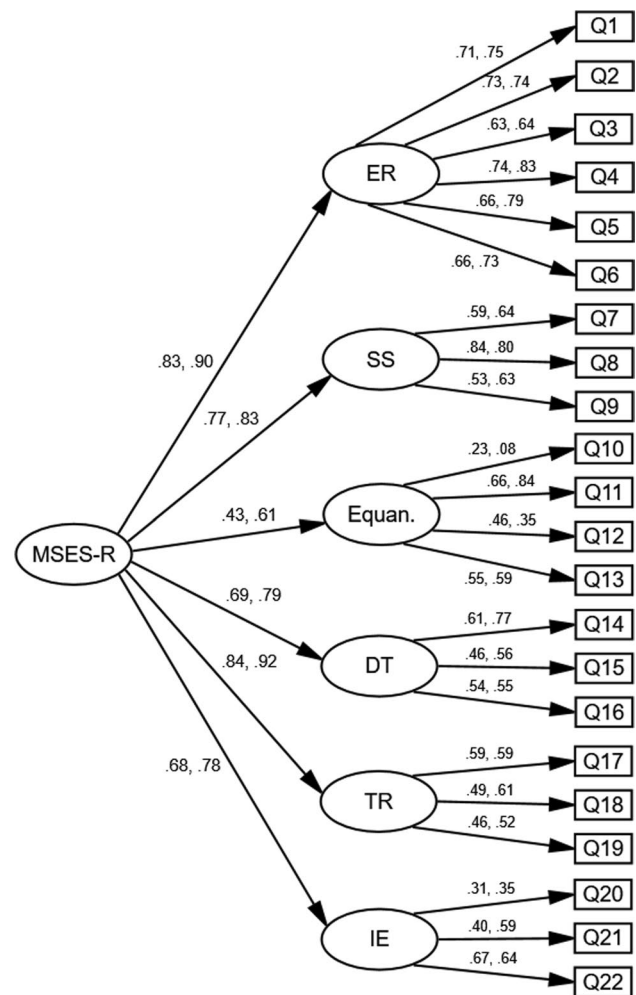

Fig. 1 CFA model showing results for the Australian clinical and community samples. Note: ER = Emotion Regulation, SS = Social Skills, Equan. = Equanimity, DT = Distress Tolerance, TR = Taking Responsibility, IE = Interpersonal Effectiveness. All values are significant at $p < .001$

Fig. 2 CFA model showing results for the Canadian clinical and community samples. Note: ER = Emotion Regulation, SS = Social Skills, Equan. = Equanimity, DT = Distress Tolerance, TR = Taking Responsibility, IE = Interpersonal Effectiveness. All values are significant at $p < .001$, except for Q10 which did not load significantly on the Equanimity factor in the community sample

Table 3 Results of multiple group CFAs to determine measurement invariance

	$\chi^2(df)$	CFI	TLI	SRMR	RMSEA	$\Delta\chi^2(df)$	ΔCFI	ΔSRMR	ΔRMSEA
Australian clinical and community samples									
Configural model	(406) = 2037.202***	.975	.971	.047	.044 (.042–.045)				
Metric model	(427) = 2469.956***	.968	.966	.051	.047 (.046–.049)	(21) = 432.754****	− .007	.004	.003
Scalar model	(442) = 2548.489***	.967	.966	.049	.047 (.046–.049)	(15) = 78.533***	− .001	− .002	.000
Canadian clinical and community samples									
Configural model	(406) = 560.356***	.988	.986	.052	.029 (.023–.34)				
Metric model	(427) = 716.785***	.977	.975	.058	.039 (.034–.043)	(21) = 156.429***	− .011	.006	.010
Scalar model	(442) = 732.074***	.977	.976	.056	.038 (.033–.043)	(15) = 15.289	.000	− .002	− .001
Australian and Canadian clinical samples									
Configural model	(406) = 1158.663***	.966	.961	.051	.043 (.040–.046)				
Metric model	(427) = 1372.707***	.957	.954	.055	.047 (.045–.050)	(21) = 214.044***	− .009	.004	.004
Scalar model	(442) = 1398.852***	.957	.955	.047	.047 (.044–.050)	(15) = 26.145*	.000	− .008	.000
Australian and Canadian community samples									
Configural model	(406) = 1438.895***	.981	.979	.046	.040 (.038–.042)				
Metric model	(427) = 1524.011***	.980	.978	.047	.040 (.038–.042)	(21) = 85.116***	− .001	− .001	.000
Scalar model	(442) = 1540.389***	.980	.979	.045	.040 (.037–.042)	(15) = 16.37	.000	− .002	.000

*** $p < .001$; * $p < .05$

Table 4 Reliability of the MSES-R overall and each facet in the Australian, Canadian, and student samples

	Australian clinical	Australian community	Canadian clinical	Canadian community
MSES-R overall	.85	.89	.83	.90
Emotion Regulation	.83	.88	.84	.89
Social Skills	.73	.72	.70	.74
Equanimity	.48	.58	.54	.53
Distress Tolerance	.58	.65	.54	.67
Taking Responsibility	.52	.60	.53	.60
Interpersonal Effectiveness	.49	.47	.45	.54

MSES-R Mindfulness-based Self-Efficacy Scale-Revised

Reliability

As shown in Table 4, the overall MSES-R measure was highly reliable across all samples, as was the Emotion Regulation facet. The reliability of the Social Skills facet was in the low to moderate range, with the reliability of the other facets often being unacceptably low. Due to the low reliability of these facets, users should rely on the total MSES-R score. Accordingly, results for the overall MSES-R will be the focus of subsequent analyses. Nonetheless, we provide the results at the facet level in the supplementary materials.

Group Differences

A 2×2 ANOVA comparing each subsample on overall MSES-R scores provided a significant main effect of sample type, differentiating clinical from community samples ($F(1, 5156) = 761.25$, $p < 0.001$, $\eta^2 = 0.13$). The Australian clinical sample ($M = 40.92$, $SD = 12.07$) and Canadian clinical sample ($M = 39.38$, $SD = 12.12$) scored significantly lower on the MSES-R than the Australian community ($M = 54.36$, $SD = 12.83$) and Canadian community ($M = 52.53$, $SD = 13.31$) samples. In addition, the main effect for country was significant ($F(1, 5156) = 12.18$, $p < 0.001$, $\eta^2 = 0.002$), indicating that the Australian sample scored higher overall on the MSES-R than the Canadian sample, although this effect was extremely weak. No significant interaction was detected.

Discussion

In line with the theoretical background and qualitative framework provided by Cayoun (2011), and the findings of Kasselis (2011) and Atalay et al. (2017), the six-factor MSES-R appeared to provide adequate model fit in all tested samples. However, while the model fit appeared to be adequate, reliability of several facets was low. Consequently, it is recommended that individual facets should not be used in isolation. Accordingly, the facet-level findings have not

been a focus here and are presented in the supplementary materials.

The results for measurement invariance suggest that the measures were not invariant between the Australian and Canadian clinical and community samples on the basis of differences in χ^2. As this is sensitive to sample size (Cheung & Rensvold, 2002), if the criteria suggested by Chen (2007) are prioritized, the observed differences in CFI, SRMR, and RMSEA generally indicated invariance. The exception to this was that there was no clear evidence that the MSES-R was invariant between the Canadian clinical and community samples. However, on the assumption that the measure was at least somewhat invariant across the Australian and Canadian clinical and community samples, the reported differences appeared to be meaningful. Notably, the community samples reported higher overall MSES-R scores than the clinical samples in both the Australian and Canadian samples. This is in line with previous research showing that the MSES-R distinguishes clinical from non-clinical samples prior to an MBI (e.g., Goldstein et al., 2018).

Study 2

Method

Participants

The sample ($N = 521$) comprised 130 males, 390 females, and one person who chose to not record their gender. From these, 100 participants also participated in the retest phase of the study, which took place 2 weeks later. Respondents were asked to record their age by selecting the appropriate range: 18–20 ($n = 109$), 21–25 ($n = 130$), 26–35 ($n = 89$), 36–45 ($n = 79$), 46–55 ($n = 54$), 56–65 ($n = 44$), 66+ ($n = 16$).

Procedure

Respondents were recruited through advertising at a metropolitan Australian university and peer referral. Respondents included undergraduate and postgraduate students. First year psychology students were provided with course credit, and no other incentives were offered to participants. Each participant indicated their informed consent through their participation. All participants received either a hardcopy or online questionnaire package containing an information sheet. All participants had the option of completing a second questionnaire package containing only the MSES, 2 weeks after completing the initial questionnaire to assess test–retest reliability.

Measures

The sample originally completed the 35-item MSES (Kasselis, 2011); however, only the 22-items comprising the MSES-R are considered here.

Depression Anxiety Stress Scale The 21-item Depression Anxiety Stress Scale (DASS-21) (Lovibond & Lovibond, 1995) includes three subscales (i.e., Depression, Anxiety, Stress), with each comprising seven items (e.g., "I found it hard to wind down"). Items are rated on a scale from 0 (did not apply to me at all) to 3 (applied to me very much, or most of the time) based on how respondents felt over the past week. Lovibond and Lovibond (1995) found each subscale to be highly reliable, and valid on the basis of correlations with existing measures of depression and anxiety. In the current study, DASS was highly reliable overall ($\omega = 0.93$), as were the Depression ($\omega = 0.89$), Anxiety ($\omega = 0.81$), and Stress ($\omega = 0.87$) subscales.

Five Facet Mindfulness Questionnaire The Five Facet Mindfulness Questionnaire (FFMQ) (Baer et al., 2006) includes 39 items across five factors. These are Nonreactivity to inner experience (e.g., "I perceive my feelings without having to react to them"), Observing (e.g., "I notice the smells and aromas of things"), Acting with awareness (e.g., "I find myself doing things without really paying attention"), Describing (e.g., "I'm good at finding the words to describe my feelings"), and Nonjudgment of experience (e.g., "I make judgments about whether my thoughts are good or bad"). Each item is rated on a 5-point scale (1 = never or very rarely true, 5 = very often or always true). Baer et al. (2006) reported the reliability of the FFMQ and each facet to be acceptable. A meta-analysis of the FFMQ found that higher overall scores and higher scores on each facet, with the exception of Observe, were consistently associated with reduced affective symptoms (Carpenter et al., 2019). In the current study, the FFMQ was highly reliable overall ($\omega = 0.89$), as were each factor: Observe ($\omega = 0.74$), Describe ($\omega = 0.73$), Acting with Awareness ($\omega = 0.81$), Nonjudgement ($\omega = 0.90$), and Nonreactivity ($\omega = 0.77$).

Previous Exposure to Meditative or Contemplative Practices We were interested to examine whether the MSES-R could differentiate participants who cultivate mindfulness and insight through meditative/contemplative practices, even to a small degree, from those who do not. Accordingly, respondents were asked if they had practiced yoga, meditation using mantras, meditation using visualization, mindfulness meditation, or record any other form of meditative or contemplative practice. They were also asked how long they had engaged in any of these practices and if they

still practiced it, how much time they dedicate to it each week. Based on the responses, two categories were created: "Meditators" ($n = 229$), who reported engaging in at least one practice for at least 15 min per week for the past 12 months, and "Non-meditators" ($n = 292$), who either had no experience with any practice or had attempted a practice fewer than six times in total and discontinued. These criteria were guided by one of the authors, who has had extensive experience in teaching insight meditation regularly in a wide range of community and clinical populations for the past 30 years. During this period, it was repeatedly shown that even a small amount of regular practice can enhance one's sense of self-efficacy. Recent studies have used a similar approach (e.g., Nyklíček, 2020; Wu et al., 2019).

Data Analyses

To investigate the model fit of the MSES-R in a separate sample, the 22-item measure was examined using CFA with the same criteria as study 1. To examine differences between meditators and non-meditators, independent samples t-tests were conducted. Prior to examining differences between these two groups, the reliability of the MSES-R and each facet using McDonald's ω was considered. Bivariate correlations were also conducted to assess the stability of the measure over time by comparing the test and retest scores, and the relationship between the MSES-R and both the FFMQ and DASS.

Fig. 3 CFA model showing the results in the student sample. Note: ER = Emotion Regulation, SS = Social Skills, Equan. = Equanimity, DT = Distress Tolerance, TR = Taking Responsibility, IE = Interpersonal Effectiveness. All values are significant at $p < .001$

Results

Model Fit

The six-factor model was found to be an excellent fit with the data, $\chi^2(203) = 238.11$, $p < 0.05$, CFI = 0.995, TLI = 0.994, SRMR = 0.045, and RMSEA = 0.018 (90% confidence interval = 0.003–0.027). The model including factor loadings is shown in Fig. 3.

Reliability

Prior to assessing differences between groups, the reliability of the MSES-R and each facet were examined. Similar to study 1, MSES-R was found to be highly reliable (ω = 0.87). With the exception of the Emotion Regulation facet (ω = 0.86), reliability of the facets was poor (ω range = 0.54–0.69). Subsequent analyses will therefore focus on the overall MSES-R, with facet-level results provided in supplementary materials.
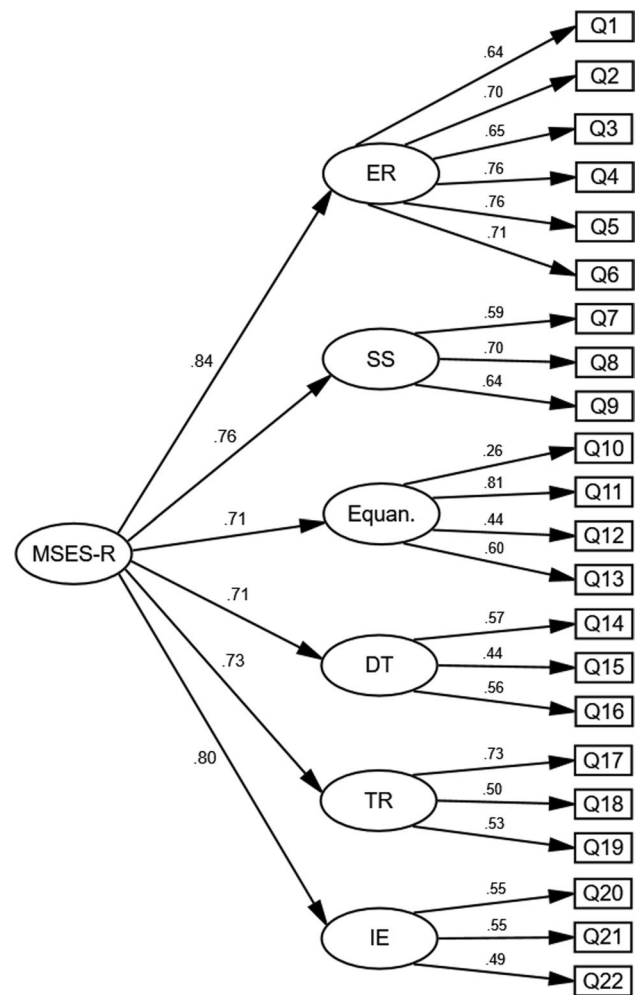
Differences Between Meditators and Non-meditators

The independent samples t-test on the MSES-R indicated that meditators ($M = 62.05$, $SD = 11.64$) scored significantly higher ($t(519) = 4.75$, $p < 0.001$, Cohen's $d = 0.42$) on the MSES-R than non-meditators ($M = 57.04$, $SD = 12.17$).

Relationships with FFMQ and DASS 21

Bivariate correlations indicated that higher overall scores on the MSES-R were significantly associated with higher scores on the FFMQ overall ($r = 0.62$, $p < 0.001$), as well as higher scores on the factors of Observe ($r = 0.23$, $p < 0.001$), Describe ($r = 0.40$, $p < 0.001$), Acting with Awareness ($r = 0.49$, $p < 0.001$), Nonjudgement ($r = 0.47$, $p < 0.001$), and Nonreactivity ($r = 0.51$, $p < 0.001$). Higher scores on

the MSES-R were also significantly associated with lower overall DASS scores ($r = -0.68$, $p < 0.001$), and on each subscale: Depression ($r = -0.60$, $p < 0.001$), Anxiety ($r = -0.56$, $p < 0.001$), and Stress ($r = -0.62$, $p < 0.001$).

Test–Retest Reliability

The results indicated a strong positive correlation between initial test scores and retest scores ($M = 52.78$, $SD = 16.06$; $M = 51.71$, $SD = 15.20$, respectively), $r = 0.88$, $n = 100$, $p < 0.01$. The shared variance was 78%.

Discussion

In the student sample, respondents who regularly engaged in meditative or contemplative practices reported higher MSES-R scores than non-meditators. Higher MSES-R scores were associated with lower self-reported symptoms of depression, anxiety, and stress, as measured by the DASS 21. This shows that the MSES-R can discriminate between meditators and non-meditators. However, these differences were modest. This may be explained by the criteria used for group allocation.

The findings also indicate that the MSES-R is reasonably stable over a limited period, as reflected by the high correlation between scores taken after a two-week interval and a shared variance of 78%. Regarding convergent validity, the MSES-R was associated with higher levels of mindfulness, as measured by the FFMQ, which corresponds to the findings using the MAAS reported by Goldstein et al. (2018). These results suggest that the MSES-R may be an effective measure when investigating the effectiveness of mindfulness-based practice or interventions in clinical or community samples.

Study 3

Method

Participants

Sixty-eight participants with existing clinical diagnoses were referred by medical doctors, psychiatrist, and other local mental health professional. All completed a 10-week Mindfulness-integrated Cognitive Behavior Therapy (MiCBT; Cayoun, 2015; Cayoun et al., 2019) program. Due to ethical constraints, no demographic data were available for these respondents. However, the sample was Australian, approximately 70% female, all respondents were aged over 18, and covered a wide age spectrum. Respondents presented with a range of diagnoses, as summarized in Table 5.

Procedure

Data were collected using pen-and-paper surveys at the start of the program, in the fifth week, and at the end of the program. All participants undertook the MiCBT program in ten individual therapy sessions conducted by the first author, as per the standard protocol (Cayoun et al., 2019). MiCBT is a transdiagnostic approach which integrates mindfulness meditation in the Burmese Vipassana tradition of U Ba Khin and Goenka (Hart, 2007) with essential CBT skills. It consists of four stages, each involving skills aimed at addressing dysfunction in four life domains. In the first stage, participants learn to regulate attention and emotions with mindfulness meditation and the cultivation of equanimity (Rogers et al., 2020). In the second stage, these skills are used to prevent avoidant behavior which reinforces or maintains the symptoms. In stage 3, skills developed in the previous stages are applied to interpersonal contexts to prevent conflict avoidance and improve assertiveness. In stage 4, participants train in developing compassion for themselves and others for the purpose of preventing relapse, by learning to combine loving-kindness meditation with ethical intentions and actions.

Table 5 Participant characteristics

Characteristic	Number (%) of participants
Primary diagnosis	
Anxiety disorder	22 (32.35%)
Adjustment disorder	12 (17.65%)
Major depressive disorder	7 (10.29%)
Major depressive episode	4 (5.88%)
Persistent depressive disorder	4 (5.88%)
Bipolar disorder	4 (5.88%)
Panic disorder	3 (4.41%)
Post-traumatic stress disorder	3 (4.41%)
Obsessive compulsive disorder	2 (2.94%)
Alcohol abuse	2 (2.94%)
Attention-deficit hyperactivity disorder	1 (1.47%)
Borderline personality disorder	1 (1.47%)
Bereavement	1 (1.47%)
Burnout	1 (1.47%)
Separation anxiety	1 (1.47%)
Number of comorbid diagnoses	
1	19 (27.94%)
2	28 (41.18%)
3	19 (27.94%)
4	2 (2.94%)

Measures

The sample completed the 22-item MSES-R, in addition to the 21-item DASS-21 (Lovibond & Lovibond, 1995). Unfortunately, due to the destruction of the pen-and-paper surveys in accordance with the necessary ethical approval, data only exists for the overall scores for the measures used. It is therefore not possible to calculate reliability coefficients. Based on the results from studies 1 and 2, the focus remained on overall scores for the MSES-R and the DASS.

Data Analyses

Bivariate correlations were conducted to examine the relationships between overall scores on the MSES-R and DASS. Repeated measures ANOVAs were conducted to examine changes across the three time points (start, 5 weeks, 10 weeks) of the MiCBT programs. This was to examine changes in scores on the MSES-R and the DASS. Repeated measures analyses were considered appropriate as responses from each respondent were matched for each period. All respondents completed the program and the survey at each time point.

Results

Correlations

Correlations between overall MSES-R and DASS scores at each time point are shown in Table 6. Higher MSES-R scores measured at each time point were associated with lower scores on the DASS at each corresponding time.

Sensitivity to Change

Results from separate repeated measures ANOVAs with Greenhouse–Geisser corrections due to violations of sphericity (i.e., Mauchly's test of sphericity $p < 0.05$) indicated a significant change in MSES-R scores ($F(1.67, 111.72) = 77.94$, $p < 0.001$, $\eta^2 = 0.54$) and DASS scores

($F(1.57, 105.49) = 170.23$, $p < 0.001$, $\eta^2 = 0.72$) over the duration of the 10-week MiCBT program. Planned contrasts indicated that changes at each time point for both measures were significant ($p < 0.001$ for all comparisons). Thus, MSES-R scores significantly increased from the start of the program ($M = 48.74$, $SD = 12.09$), to mid-treatment ($M = 54.94$, $SD = 10.46$), and increased further at post-treatment ($M = 65.03$, $SD = 11.64$). Conversely, DASS scores at the start of treatment ($M = 61.04$, $SD = 24.18$) decreased at mid-treatment ($M = 34.21$, $SD = 18.02$) and again at post-treatment ($M = 18.19$, $SD = 14.95$).

Discussion

As predicted, MSES-R scores increased significantly as a function of progress throughout a 10-week MiCBT program, showing that the instrument is sensitive to this kind of clinical intervention. Increases in MSES-R scores were highly correlated with a significant decrease in depression, anxiety, and stress scores on the DASS. This is in support of previous studies showing good sensitivity to change following an MBI (Alexander et al., 2012; Goldstein et al., 2018).

General Discussion

The study investigated the factor structure, measurement invariance, and reliability of the 22-item MSES-R in clinical and community samples. The study also examined group differences (i.e., community and clinical Australian and Canadian samples; meditators and non-meditators), sensitivity to change in MSES-R scores during and following therapeutic intervention, and the convergent validity of the measure in a student sample. The findings aligned with expectations, supporting a six-factor model which generally displayed invariance across clinical and community samples. In addition, the clinical sample scored lower on the MSES-R than a community sample, meditators scored higher on the MSES-R than non-meditators, and MSES-R scores were found to increase in conjunction with the completion of a mindfulness-based

Table 6 Correlations between the MSES-R and DASS at each time point

	1	2	3	4	5	6
1. MSES-R Time 1	–					
2. MSES-R Time 2	.68***	–				
3. MSES-R Time 3	.40***	.59***	–			
4. DASS Time 1	− .62***	− .52***	− .27*	–		
5. DASS Time 2	− .45***	− .66***	− .47***	.63***	–	
6. DASS Time 3	− .31***	− .50***	− .71***	.36***	.62***	–

*$p < .05$; **$p < .01$; ***$p < .001$

MSES-R Mindfulness-based Self-Efficacy Scale-Revised, *DASS* Depression Anxiety Stress Scale

intervention. Furthermore, the MSES-R displayed convergent validity by correlating with lower symptoms of depression, anxiety, and stress in community and clinical samples, and higher FFMQ scores in a community sample. Despite this, there appeared to be shortcomings in the MSES-R, including the poor reliability for most facets.

As suggested by Stanley and Edwards (2016), good model fit and poor reliability may occur together in situations where there are an insufficient number of items on factor(s). In addition, the approach taken to CFA in the current study (see Sellbom & Tellegen, 2019) treated the MSES-R items as ordinal rather than continuous. As argued by Sellbom and Tellegen (2019), in the past, many models may have been rejected when they should not have been, as items have been treated as continuous measures when this was not the case. Hence, model fit may have appeared to have been poor in the current study if the items were treated as continuous variables. While the current results support the MSES-R as comprising six facets, if future research intends to focus on facets independently, additional items must be added.

The measure may also benefit from revising the wording of some items. For example, it may be that items such as "I find it difficult to make new friends" or "I can resolve problems easily with my partner (or best friend if single)" may be influenced by traits such as introversion or shyness, or are too specific to certain types of relationships. Such issues may have contributed to variability in responses and thus a reduction in reliability coefficients. Despite these issues, the overall MSES-R was reliable and appeared to be valid and produced meaningful differences between groups.

It must be noted, however, that in the second study, the differences between those engaged in meditative or contemplative practices and those who are not were modest. This may be explained by the criteria used for group allocation. As we were interested in examining the MSES-R's ability to discriminate between meditators and non-meditators, even to a small degree, the criteria for including participants in the meditators' group were widely inclusive and not limited to mindfulness meditators. Those who practiced yoga and other contemplative practices, including meditations using mantras or visualization, were included in the meditators' group if they regularly engaged in their practice for at least 15 min per week for the past 12 months. It is not clear whether the MSES-R performs similarly with mindfulness meditators and meditators of other methods. It is likely that not all meditations or yoga methods equally improve one's perceived self-efficacy in coping with daily stressors in a mindful and equanimous way. Moreover, although 15-min practice per week can indeed make a difference on one's ability to improve emotion processing (Wu et al., 2019), the required amount of practice to develop a sense of self-efficacy detectable by the MSES-R is unclear. Hence, variations in the

quality and amount of meditative practice in the meditators' group may account for the observed effect sizes.

In the third study, the ability to show both temporal stability and sensitivity to change in mindfulness-based self-efficacy following an MBI is an important advantage in clinical contexts. This is not always evident in some mindfulness-related measures. For example, a meta-analytic study examined the reliability of brief mindfulness training in reducing negative affectivity in 65 randomized controlled trials (RCTs), including 5,489 participants predominantly without experience in meditation (Schumer et al., 2018). While the results showed a small but significant effect of brief MBIs on reducing negative affect, no significant differences in effect size were found between clinical and non-clinical samples. As it was not clear whether the lack of group differences was affected by the type of measures used in these studies, the authors suggested using measures other than just trait-mindfulness, such as ones that may be important for determining the likelihood of benefits.

According to self-efficacy theory (Bandura, 1997), the two key predictors of behavior are the individual's perceived self-efficacy and outcome expectancies. Given the evidence that self-efficacy beliefs can play an important role in determining successful outcomes (Bandura, 2006), we suggest that measuring mindfulness-related self-efficacy is of benefit to the field. Particularly in the clinical context, other authors have also suggested that "Clinicians administering mindfulness-based interventions should be aware of the role of coping self-efficacy in the relationship between mindfulness and emotion regulation" (Luberto et al., 2013, p. 274). Based on the present results and the findings of previous studies, the MSES-R was found to contribute to the field of mindfulness research by providing another dimension of behavior change facilitated by mindfulness training.

Limitations and Directions for Future Research

The present study was limited in several ways. For instance, in study 2, the criteria to allocate participants to a meditators group may have been too inclusive to allow greater examination of the extent to which the MSES-R can separate meditators from non-meditators. Future research will need to distinguish between different types of practice more comprehensively. For example, someone who has meditated for 30 min daily for several years may have attained very different insights and perspectives than someone who has used a mindfulness phone-based app for 15 min a week over the course of 6 months. Future studies could explore the extent to which the MSES-R can differentiate between meditators and non-meditators by varying the respondents' experience and mindfulness practice history. We predict that MSES-R scores will be greater for more experienced mindfulness meditators. Additionally, insights associated

with mindfulness, such as nonattachment, may arise through experiencing difficult life circumstances in conjunction with non-meditative practices such as formal psychotherapy (Whitehead et al., 2018). Future studies could compare the MSES-R scores of mindfulness meditators with those of respondents who undertook a course of unrelated therapy.

A further limitation of the current study is the possibility of a selection bias. The Australian and Canadian clinical and community samples comprised individuals who freely chose to access a website and completed the MSES-R. The student sample was similarly affected by self-selection and peer-referral. It could reasonably be considered that respondents in these samples had an existing interest in mindfulness. Thus, the current results may not hold in samples where there is no interest in or active avoidance of mindfulness. The cultural generalizability of the current findings is also limited by the focus on Australian and Canadian samples. However, the current findings in support of the model fit and validity of the MSES-R correspond with those from a Turkish adaptation of the measure (Atalay et al., 2017). Given the expanding use of mindfulness measures across cultures (e.g., Lopez-Maya et al., 2019), cross-cultural consistency studies of the MSES-R are needed.

Furthermore, Australian and Canadian clinical samples were based on respondents self-reporting current and past clinical symptoms, therapy attendance, and clinical diagnoses provided by their therapists. It was therefore not possible to verify if these participants did genuinely meet respective diagnostic criteria. Future studies should attempt to address this by obtaining verifiable information on clinical diagnosis. This would also permit an analysis of possible differences in MSES-R responses based on diagnosis. Currently, there is no way of knowing if individuals with different mental health conditions would relate to MSES-R items differently. Moreover, in study 3, the absence of item-by-item analysis did not permit the calculation of reliability coefficients. Future studies using clinical samples will need to avoid this limitation. Finally, participants in the clinical intervention study were given the questionnaires, including the DASS-21 and MSES-R by the therapist. It cannot be ruled out that participants unintentionally reported inflated scores for socially desirable reasons. Future studies could control for social desirability.

Author Contributions B. C.: co-designed the study, collected online data, analyzed the data, collaborated with the design, and writing and editing of the manuscript. B. E.: co-designed the study, performed the overall statistical analyses, and collaborated with the writing of the study. N. K.: collected and analyzed the data of the student sample. G. B.: collaborated with online data collection and analysis. C. S.: supervised the study and collaborated with early stages of data analyses. All authors approved the final version of the manuscript for submission.

Declarations

Informed Consent The study was approved by the Human Research Ethics Committee of the University of Tasmania, Australia (HREC Ref H11163). Participants were fully informed of the procedure and aims of the study and gave their informed consent.

Conflict of Interest The authors declare no competing interests.

References

Alexander, V., Tatum, B. C., Auth, C., Takos, D., Whittemore, S., & Fidaleo, R. (2012). A study of mindfulness practices and cognitive therapy: Effects on depression and self-efficacy. *International Journal of Psychology and Counselling, 4*(9), 115–122. https://doi.org/10.5897/IJPC12.030

Atalay, Z., Aydin, U., Bulgan, G., Taylan, R. D., & Özgülük, S. B. (2017). Mindfulness-based Self Efficacy Scale-Revised (MSES-R): Turkish adaptation study. *Elementary Education Online, 16*(4), 1803–1815.

Australian Bureau of Statistics. (2021). *National, state and territory population, Dec 2020*. Commonwealth of Australia. Retrieved August 4, 2021, from https://www.abs.gov.au/statistics/people/population/national-state-and-territory-population/dec-2020#abs-stat-datasets

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

Bandura, A. (1997). *Self-efficacy: The exercise of control*. Freeman.

Bandura, A. (2006). Guide for constructing self-efficacy scales. In F. Pajares & T. Urdan (Eds.), *Self-efficacy beliefs of adolescents* (Vol. 5, pp. 307–337). Information Age Publishing.

Bednar, K., Voracek, M., & Tran, U. S. (2020). Common factors underlying the five facets of mindfulness and proposed mechanisms: A psychometric study among meditators and non-meditators. *Mindfulness, 11*, 2804–2817. https://doi.org/10.1007/s12671-020-01492-6

Bergomi, C., Ströhle, G., Michalak, J., Funke, F., & Berking, M. (2012). Facing the dreaded: Does mindfulness facilitate coping with distressing experiences? A moderator analysis. *Cognitive Behaviour Therapy, 42*(1), 21–30. https://doi.org/10.1080/16506073.2012.713391

Bishop, S. R., Lau, M., Shapiro, S., Carlson, L., Anderson, N. D., Carmody, J., Segal, Z. V., Abbey, S., Speca, M., Velting, D., & Devins, G. (2004). Mindfulness: A proposed operational definition. *Clinical Psychology: Science and Practice, 11*(3), 230–241. https://doi.org/10.1093/clipsy.bph077

Block-Lerner, J., Salters-Pedneault, K., & Tull, M. T. (2005). Assessing mindfulness and experiential acceptance: Attempts to capture inherently elusive phenomena. In S. M. Orsillo & L. Roemer (Eds.), *Acceptance and mindfulness-based approaches to anxiety* (pp. 71–99). Springer.

Bodhi, B. (2005). *In the Buddha's words: An anthology of discourses from the Pali Canon*. Wisdom Publications.

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

Buddhaghosa, H. (2010). *The Path of Purification (Visuddhimagga): The classic manual of Buddhist doctrine and meditation*, trans. Nyanamoli Himi. Buddhist Publication Society.

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

Cayoun, B. A. (2011). *Mindfulness-integrated CBT: Principles and practice*. Wiley.

Cayoun, B. A. (2015). *Mindfulness-integrated CBT for well-being and personal growth: Four steps to enhance inner, self-confidence and relationships*. Wiley.

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.

Chen, F. F. (2007). Sensitivity of goodness of fit indexes to lack of measurement invariance. *Structural Equation Modelling, 14*(3), 464–504. https://doi.org/10.1080/10705510701301834

Chang, V. Y., Palesh, O., Caldwell, R., Glasgow, N., Abramson, M., Luskin, F., et al. (2004). The effects of a mindfulness-based stress reduction program on stress, mindfulness self-efficacy, and positive states of mind. *Stress and Health: Journal of the International Society for the Investigation of Stress, 20*(3), 141–147. https://doi.org/10.1002/smi.1011

Cheung, G. W., & Rensvold, R. B. (2002). Evaluating goodness-of-fit indexes for testing measurement invariance. *Structural Equation Modeling, 9*(2), 233–255. https://doi.org/10.1207/S15328007SEM0902_5

Delacre, M., Leys, C., Mora, Y. L., & Lakens, D. (2019). Taking parametric assumptions seriously: Arguments for the use of Welch's F-test instead of the classical F-test in one-way ANOVA. *International Review of Social Psychology, 32*(1), 13. https://doi.org/10.5334/irsp.198

Desbordes, G., Gard, T., Hoge, E. A., Hölzel, B. K., Kerr, C., Lazar, S. W., Olendzki, A., & Vago, D. R. (2015). Moving beyond mindfulness: Defining equanimity as an outcome measure in meditation and contemplative research. *Mindfulness, 6*(2), 356–372. https://doi.org/10.1007/s12671-013-0269-8

Eberth, J., Sedlmeier, P., & Schafer, T. (2019). PROMISE: A model of insight and equanimity as the key effects of mindfulness meditation. *Frontiers in Psychology, 10*, 2389. https://doi.org/10.3389/fpsyg.2019.02389

Fabrigar, L. R., Wegener, D. T., MacCallum, R. C., & Strahan, E. J. (1999). Evaluating the use of exploratory factor analysis in psychological research. *Psychological Methods, 4*(3), 272–299. https://doi.org/10.1037/1082-989X.4.3.272

Feng, X. J., Krägeloh, C. U., Billington, D. R., & Siegert, R. J. (2018). To what extent is mindfulness as presented in commonly used mindfulness questionnaires different from how it is conceptualized by senior ordained Buddhists? *Mindfulness, 9*(2), 441–460. https://doi.org/10.1007/s12671-017-0788-9

Goldstein, E., Topitzes, J., Brown, R. L., & Barrett, B. (2018). Mediational pathways of meditation and exercise on mental health and perceived stress: A randomized controlled trial. *Journal of Health Psychology, 25*(12), 1816–1830. https://doi.org/10.1177/1359105318772608

Grabovac, A. (2015). The stages of insight: Clinical relevance for mindfulness-based interventions. *Mindfulness, 6*, 589–600. https://doi.org/10.1007/s12671-014-0294-2

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

Gu, J., Strauss, C., Bond, R., & Cavanagh, K. (2015). How do mindfulness-based cognitive therapy and mindfulness-based stress reduction improve mental health and wellbeing? A systematic review and meta-analysis of mediation studies. *Clinical Psychology Review, 37*, 1–12. https://doi.org/10.1016/j.cpr.2015.01.006

Hart, W. (1987). *The art of living: Vipassana meditation as taught by S. N. Goenka*. Harper & Row.

Hu, L., & Bentler, P. M. (1999). Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. *Structural Equation Modelling, 6*, 1–55. https://doi.org/10.1080/10705519909540118

Hu, L. T., Bentler, P. M., & Kano, Y. (1992). Can test statistics in covariance structure analysis be trusted? *Psychological Bulletin, 112*(2), 351–362. https://doi.org/10.1037/0033-2909.112.2.351

JASP Team (2020). JASP (Version 0.14.1) [Computer software].

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

Kabat-Zinn, J. (2006). *Coming to our senses: Healing ourselves and the world through mindfulness*. Piatkus.

Kasselis, N. (2011). *The Mindfulness-based Self-Efficacy Scale-Revised*. Unpublished dissertation. University of Tasmania, Australia.

Li, C.-H. (2016). The performance of ML, DWLS, and ULS estimation with robust corrections in structural equation models with ordinal variables. *Psychological Methods, 21*(3), 369–387. https://doi.org/10.1037/met0000093

Lopez-Maya E., Olmstead, R., & Irwin, M. R. (2019). Mindfulness meditation and improvement in depressive symptoms among Spanish- and English speaking adults: A randomized, controlled, comparative efficacy trial. PLoS ONE 14(7): e0219425. https://doi.org/10.1371/journal.pone.0219425

Lovibond, S. H., & Lovibond, P. F. (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

Luberto, C., Cotton, S., McLeish, A. C., Mingione, C. A., & O'Bryan, E. M. (2013). Mindfulness skills and emotion regulation: The mediating role of coping self-efficacy. *Mindfulness, 5*, 373–380. https://doi.org/10.1007/s12671-012-0190-6

Nyklíček, I. (2020). Aspects of self-awareness in meditators and meditation-naïve participants: Self-report versus task performance. *Mindfulness, 11*, 1028–1037. https://doi.org/10.1007/s12671-020-01318-5

McNeish, D. (2018). Thanks coefficient alpha, we'll take it from here. *Psychological Methods, 23*(3), 412–433. https://doi.org/10.1037/met0000144

Rogers, H., Shires, A. G., & Cayoun, B. A. (2020). Development and validation of the Equanimity Scale-16. *Mindfulness, 12*, 107–120. https://doi.org/10.1007/s12671-020-01503-6

Schumer, M. C., Lindsay, E. K., & David Creswell, J. (2018). Brief mindfulness training for negative affectivity: A systematic review and meta-analysis. *Journal of Consulting and Clinical Psychology, 86*(7), 569–583. https://doi.org/10.1037/ccp0000324

Schwarzer, R. (2008). Modeling health behavior change: How to predict and modify the adoption and maintenance of health behaviors. *Applied Psychology, 57*(1), 1–29. https://doi.org/10.1111/j.1464-0597.2007.00325.x

Segal Z. V., Williams J. M. G., & Teasdale J. D. (2002). Mindfulness-based cognitive therapy for depression: A new approach to preventing relapse. New York: Guilford

Sellbom, M., & Tellegen, A. (2019). Factor analysis in psychological assessment research: Common pitfalls and recommendations. *Psychological Assessment, 31*(12), 1428–1441. https://doi.org/10.1037/pas0000623

Shi, D., & Maydeu-Olivares, A. (2020). The effect of estimation methods on SEM fit indices. *Educational and Psychological Measurement, 80*(3), 421–445. https://doi.org/10.1177/0013164419885164

Stanley, L. M., & Edwards, M. C. (2016). Reliability and model fit. *Educational and Psychological Measurement, 76*(6), 976–985. https://doi.org/10.1177/0013164416638900

Statistics Canada (2019). *Annual demographic estimates: Canada, provinces and territories, 2018*. Statistics Canada catalogue. Retrieved August 4, 2021, from https://www150.statcan.gc.ca/n1/pub/91-215-x/2018002/sec2-eng.htm

Van Dam, N. T., Earlywine, M., & Borders, A. (2010). Measuring mindfulness? An item response theory analysis of the Mindfulness Attention Awareness Scale. *Personality and Individual Differences, 49*(7), 805–810. https://doi.org/10.1016/j.paid.2010.07.020

Van Dam, N. T., Earlywine, M., & Danoff-Burg, S. (2009). Differential item function across meditators and non-meditators on the Five Facet Mindfulness Questionnaire. *Personality and Individual Differences, 47*(5), 516–521. https://doi.org/10.1016/j.paid.2009.05.005

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

Van de Schoot, R., Lugtig, P., & Hox, J. (2012). A checklist for testing measurement invariance. *European Journal of Developmental Psychology, 9*(4), 486–492. https://doi.org/10.1080/17405629.2012.686740

Van Dijk, S., Jeffrey, J., & Katz, M. R. (2013). A randomized, controlled, pilot study of dialectical behavior therapy skills in a psychoeducational group for individuals with bipolar disorder. *Journal of Affective Disorders, 145*(3), 386–393. https://doi.org/10.1016/j.jad.2012.05.054

Voith, L. A., Russell, K., Lee, H., & Anderson, R. E. (2020). Adverse childhood experiences, trauma symptoms, mindfulness, and intimate partner violence: Therapeutic implications for marginalized men. *Family Process, 12533*,. https://doi.org/10.1111/famp.12533

Whitehead, R., Bates, G., & Elphinstone, B. (2018). Stories of suffering and growth: An investigation of the lived experience of nonattachment. *Contemporary Buddhism, 19*(2), 448–475. https://doi.org/10.1080/14639947.2018.1572311

Wu, R., Liu, L. L., Zhu, H., Su, W. J., Cao, Z. Y., Zhong, S. Y., Liu, X. H., & Jiang, C. L. (2019). Brief mindfulness meditation improves emotion processing. *Frontiers in Neuroscience, 13*, 1074. https://doi.org/10.3389/fnins.2019.01074

Xia, Y., & Yang, Y. (2019). RMSEA, CFI, and TLI in structural equation modeling with ordered categorical data: The story they tell depends on the estimation methods. *Behavior Research Methods, 51*(1), 409–428. https://doi.org/10.3758/s13428-018-1055-2

Zhang, D., Lee, E. K. P., Mak, E. C. W., Ho, C. Y., & Wong, S. Y. S. (2021). Mindfulness-based interventions: An overall review. *British Medical Bulletin, 138*(1), 41–57. https://doi.org/10.1093/bmb/ldab005

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