



Published in final edited form as:

J Couns Psychol. 2020 January ; 67(1): 115–122. doi:10.1037/cou0000369.

Does practice quality mediate the relationship between practice time and outcome in mindfulness-based stress reduction?

Simon B. Goldberg,

University of Wisconsin – Madison

Cara Knoeppel,

University of Wisconsin – Madison and Florida State University

Richard J. Davidson,

University of Wisconsin – Madison

Lisa Flook

University of Wisconsin – Madison and Learning Policy Institute

Abstract

Interventions based on mindfulness meditation are increasingly common and evidence exists supporting their use. However, questions remain regarding treatment mechanisms accounting for beneficial effects. The current study examined one candidate mechanism - mindfulness practice quality - as a mediator of the link between practice time and outcome within mindfulness-based stress reduction (MBSR). Participants ($n=96$) completed measures of mindfulness and psychological symptoms at baseline and post-treatment. A weekly questionnaire assessed practice time and quality over the eight weeks of MBSR. Multilevel models accounted for nesting within participants, MBSR groups, and instructors. Results generally supported the reliability and validity of a weekly single-item practice quality measure. Greater practice time was associated with improved practice quality ($r=.48$). Increases in practice quality predicted improvements in self-report mindfulness and psychological symptoms ($\beta=.35, .30, \text{ and } -.19, p<.05,$), but not behavioral mindfulness ($\beta=-.02, p = .879$). In multilevel mediation models, improved practice quality mediated the link between practice time and changes in self-report mindfulness, suggesting improved practice quality functions as a mechanism linking practice time and outcome in MBSR. Future research evaluating practice quality in clinical samples, in tandem with intensive sampling paradigms (e.g., experience sampling) and objective (behavioral, physiological) outcomes may be warranted.

Keywords

mindfulness; meditation; mindfulness-based stress reduction; practice quality

Correspondence concerning this article should be addressed to Simon B. Goldberg, Department of Counseling Psychology, 335 Education Building, 1000 Bascom Mall, University of Wisconsin - Madison, Madison, WI, 53706. sbgoldberg@wisc.edu.

Author Note: Simon B. Goldberg, Department of Counseling Psychology and Center for Healthy Minds, University of Wisconsin - Madison; Cara Knoeppel, Center for Healthy Minds, University of Wisconsin - Madison and Department of Educational Psychology and Learning Systems, Florida State University; Richard J. Davidson, Center for Healthy Minds, Department of Psychology, and Department of Psychiatry, University of Wisconsin - Madison; Lisa Flook, Center for Healthy Minds, University of Wisconsin - Madison and Learning Policy Institute.

“Practice does not make perfect, only perfect practice makes perfect.”

- Vince Lombardi

Psychosocial interventions based on mindfulness meditation have become an increasingly visible part of the healthcare landscape (Demarzo, Cebolla, & Garcia-Campayo, 2015). Epidemiological data suggest that among United States adults, utilization of meditation has more than tripled between 2012 and 2017 (4.1% to 14.2% used in the past year; Clarke, Barnes, Black, Stussman, & Nahin, 2018) and psychological reasons are among the most common motivations for practicing mindfulness meditation (e.g., 92% report using for stress management; Burke, Lam, Stussman, & Yang, 2017). While the original mindfulness-based intervention was designed to address chronic pain (Mindfulness-Based Stress Reduction [MBSR]; Kabat-Zinn, 2013), a variety of mindfulness-based interventions (MBIs) have now been developed targeting a variety of psychiatric conditions including depression, substance abuse, and eating disorders. Recent meta-analytic evidence suggests that these interventions perform on par with frontline, evidence-based treatments when compared directly (Goldberg et al., 2018).

The theoretical underpinnings of MBIs are derived from Buddhist contemplative practices emphasizing the cultivation of mindful attention that is present-focused and nonreactive (Kabat-Zinn, 2013). This occurs through increased meta-awareness (i.e., awareness of one's own mental processes) coupled with dereification of the contents of thought and perception (i.e., seeing thoughts as “just thoughts” rather than as objective reality; Wielgosz, Goldberg, Kral, Dunne, & Davidson, 2019). Within the theory of MBIs, mindfulness is developed through formal (e.g., sitting meditation) and informal (e.g., awareness in daily life) meditation practice, which lead to increased emotion regulation and improved well-being in the context of daily life (Wielgosz et al., 2019).

Research on MBIs has begun to turn from developing novel treatments to identifying key treatment ingredients and mechanisms of action (Gu, Strauss, Bond, & Cavanagh, 2015). These efforts are important not only for deepening scientific understanding of MBIs but also have direct relevance for optimizing the efficacy and acceptability of MBIs. Mindfulness meditation practice is a central treatment ingredient in MBIs and standardized MBIs often recommend participants engage in a substantial amount of home practice (e.g., 45 minutes per day, six days per week; Kabat-Zinn, 2013). However, overly burdensome home practice requirements may limit treatment adherence and acceptability (Adams et al., 2018). Thus, research investigating this aspect of MBIs is important at the current stage of research, both for a deeper understanding of treatment mechanisms and exploring potential facilitators and barriers to dissemination and implementation (Crane & Kuyken, 2013; Dimidjian & Segal, 2015).

While central to the theory of MBIs, mindfulness practice itself has to date remained largely a black box. Available evidence suggests that at the participant level, greater home practice is associated with larger reductions in psychological symptoms ($r = .26$, $k = 43$ studies; Parsons, Crane, Parsons, Fjorback, & Kuyken, 2017). Yet, the mechanisms through which practice time actually operates are unclear. Lacking a clear understanding of how practice

time impacts outcome may limit participants' and instructors' ability to use this time most effectively.

One possibility consistent with the theory underlying MBIs is that practice time is important insofar as it facilitates improvement in the *quality* of one's mindfulness practice (i.e., implementation of mindfulness skills during practice periods; Del Re, Fluckiger, Goldberg, & Hoyt, 2013). This notion is in keeping with literature on the development of expertise emphasizing focused, effortful practice versus simply acquiring experience or engaging in a more hobby-like or automated fashion (Ericsson, 2008). In theory, higher quality practice may both facilitate and reflect an individuals' acquisition of mindfulness skills relevant for daily life and therefore have a more proximal impact on treatment outcomes than practice time.

Two previous studies have examined the construct of practice quality within MBIs. Del Re, Fluckiger, Goldberg and Hoyt (2013) developed a seven-item measure - the Practice Quality - Mindfulness (PQM) - designed to assess MBSR participants' effortful, deliberate engagement during a period of mindfulness meditation practice. Results supported the measure's internal consistency, convergent and construct validity, and changes in PQM scores predicted pre-post changes in psychological symptoms. In a second study, Goldberg, Del Re, Hoyt and Davis (2014) examined the PQM within a mindfulness-based smoking cessation intervention. The PQM again showed internal consistency and increased over the course of training. Replicating and extending Del Re et al., Goldberg et al. found that changes in practice quality (and not practice time) predicted psychological functioning at five-month follow up.

These studies highlight the potential importance of practice quality but fail to assess directly the possibility that improvement in practice quality functions as a proximal mechanism (i.e., mediator) linking practice time and outcome. This possibility is crucial to evaluate, as it can directly inform the implementation of MBIs. If practice quality indeed mediates the relationship between practice time and outcome, instructors and participants could be encouraged to emphasize and perhaps monitor quality (in addition to quantity) routinely over the course of MBIs (as in routine outcome monitoring; Wampold, 2015). Likewise, researchers could evaluate ways of enhancing this aspect of treatment in efforts to optimize the efficacy and acceptability of MBIs. Further, the existing literature examining the PQM as a predictor of outcome has been limited by small sample sizes (e.g., $ns = 19$ and 43 ; Del Re et al., 2013; Goldberg et al., 2014) and exclusive reliance on self-report measures of psychological variables.

The current study aimed to address these limitations and directly assess improvement in practice quality as a mediator linking practice time and outcomes within MBIs. This was examined in a sample of 96 MBSR participants drawn from a community setting. We included a combination of self-report measures assessing mindfulness and psychological symptoms along with a behavioral task purported to measure mindfulness. Formal tests of mediation (Tingley, Yamamoto, Hirose, Keele, & Imai, 2014) were conducted within a multilevel model framework accounting for participant nesting within groups and instructors. Of note, the current study was not designed to test the efficacy of MBSR.

Method

Participants and procedure

Participants were recruited from a local MBSR program in a medium-sized, Midwestern city. Members of the research team attended MBSR orientation sessions and provided a verbal description of study procedures to prospective participants. All data collection occurred online. Procedures included the completion of an online survey and behavioral task (described below) at pre- and post-MBSR along with a brief, weekly questionnaire assessing mindfulness practice time and quality. Participants were paid \$50 for completing pre- and post-treatment measures. Study procedures were approved by the relevant Institutional Review Board.

A total of 215 individuals were offered participation in the study, with 122 providing consent and completing baseline measures, 96 completing at least two weekly questionnaire assessing practice time and quality (necessary for longitudinal growth curve models), and 89 completing all study procedures. Participants who completed weekly questionnaires did not differ from those who did not on any study variable ($p > .050$). Demographics of the analytical sample are presented in Supplemental Materials Table 1. The sample was predominantly white (82.3%), female (70.8%), college graduates (85.4%), aged 47.61 ($SD = 12.72$).

Following the orientation meeting, participants engaged in a standard, eight-week MBSR course (Kabat-Zinn, 2013). Participants were drawn from 18 classes taught by nine experienced instructors. Classes had on average 15.78 members ($SD = 2.13$) and 5.67 study participants ($SD = 2.35$, range = 2 to 11). Instructors taught between one and four classes (mean = 2, $SD = 1.12$). The MBSR program at this site has been offering courses for over 20 years. The MBSR instructors had on average 16.72 years of personal meditation practice ($SD = 6.82$), 12.11 years ($SD = 5.23$) of meditation teaching experience, and had taught 24 MBSR courses ($SD = 14.96$). Instructors were on average 56.25 years old ($SD = 9.04$), white (88.9%), and female (66.7%).

Measures

Baseline measures.—Two measures were collected only at baseline. These included a basic demographic questionnaire and the *Marlowe-Crowne Social Desirability Scale - Short Form* (MCSD; Reynolds, 1982). The MCSD is intended to capture participants' tendency to respond in a socially desirable (rather than accurate) way and has been used to model this response set bias. Higher scores indicate a greater tendency to respond in socially desirable ways. The 13-item MCSD which was used has shown good convergent validity with longer versions (Reynolds, 1982); internal consistency was acceptable (Kuder-Richardson = .72).

Pre- and post-treatment measures.

Mindfulness. Three self-report measures and one behavioral task assessed aspects of mindfulness (trait mindfulness, mindfulness practice quality) at pre- and post-treatment. *Five Facet Mindfulness Questionnaire* (FFMQ; Baer, Smith, Hopkins, Krietemeyer, & Toney, 2006) is a widely used 39-item measure designed to assess mindfulness across five

dimensions: acting with awareness, observing, describing, nonreactivity, and nonjudging. Responses range from 1 (never or rarely true) to 5 (very often or always true). A total score was computed across all items (range = 39 to 195) to reduce the number of tests conducted and based on prior work demonstrating a hierarchical structure including all five factors (Baer et al., 2006). Higher scores indicate higher self-report trait mindfulness. Internal consistency was high ($\alpha = .94$).

Mindful Attention Awareness Scale (MAAS; Brown & Ryan, 2003) is a widely used 15-item measure of present-moment awareness. Responses range from 1 (almost always) to 6 (almost never). The MAAS is responsive to mindfulness training, correlates in expected directions with measures of wellbeing, and differentiates between individuals with and without meditation experience (Brown & Ryan, 2003). A total score was computed across all items (range = 15 to 90), with a higher score reflecting greater mindful attention in daily life. Internal consistency reliability was high ($\alpha = .90$).

Breath counting task (BCT; Levinson, Stoll, Kindy, Merry, & Davidson, 2014) was administered as a behavioral measure of mindfulness. Based on a traditional mindfulness meditation technique, this task requires participants to count their breaths from one to nine (Levinson et al., 2014). Participants are instructed to press the “j” key for all breaths between one and eight and press the “k” key for the ninth breath. Breath counting accuracy is computed as the proportion of responses in which the appropriate key is pressed, with a higher accuracy score interpreted as greater behavioral mindfulness (range = 0.0 to 1.0). The initial study validating the BCT found the measure to be reliable across time, correlated with self-report measures of mindfulness, able to differentiate meditators from meditation naive controls, and associated with mood and meta-awareness (Levinson et al., 2014). A subsequent study also showed adequate test-retest reliability ($ICC = .48$) and found the BCT to be associated with measures of attention, although not with self-report mindfulness (Wong et al., 2018).

Practice Quality-Mindfulness (PQM; Del Re, Fluckiger, Goldberg, & Hoyt, 2013) was administered at pre- and post-treatment directly following the BCT. The PQM is a seven-item self-report measure intended to assess the degree to which an individual engages with mindfulness practice as intended (e.g., “During practice, I was struggling against having certain experiences [e.g., unpleasant thoughts, emotions, and/or bodily sensations].”) Responses range from 0 to 100%. As noted, the PQM has shown structural, convergent, construct, and predictive validity (Del Re et al., 2013; Goldberg et al., 2014). A total score was computed by averaging across all items (range = 0 to 100), with higher scores indicating higher self-report practice quality. Internal consistency reliability was adequate in the current sample ($\alpha = .71$).

Depression Anxiety and Stress Scale (DASS; Lovibond & Lovibond, 1995) is a widely used, 42-item measure assessing symptoms associated with, anxiety, and stress in the past week. The DASS has shown evidence for reliability (internal consistency, test-retest), structural, convergent, and discriminant validity and has been used to study both clinical and non-clinical manifestations of depression, anxiety, and stress (Brown, Chorpita, Korotitsch, & Barlow, 1997; Crawford & Henry, 2003). Responses range from 0 (did not apply to me at

all) to 3 (applied to me very much or most of the time). A total score was computed by summing across all items (range = 0 to 126) with a higher score indicating the presence of more psychological symptoms in the past week. Internal consistency reliability was high in the current sample ($\alpha = .93$).

Weekly questionnaire.

Practice time.: Two items assessed average daily formal (“Approximately how many minutes per day did you engage in formal mindfulness practice [e.g., sitting meditation, body scan] this week?”) and informal (“Approximately how many minutes per day did you engage in informal mindfulness practice [e.g., dropping-in, mindful eating] this week?”) practice in the past week. Total practice time was computed by summing across formal and informal practice. Practice time values were inspected for the possibility of reporting errors. Values greater than 120 minutes of either formal or informal practice were deemed likely errors (given the MBSR recommendation of 45 minutes). These responses were quite rare (1.6% and 0.5% of responses for formal and informal minutes, respectively) and were treated as missing in the primary models (i.e., participants’ average practice time was computed without them). A sensitivity analysis was also conducted retaining raw practice time values.

Single-item PQM.: In order to reduce participant burden and increase adherence to weekly online questionnaires, a single item assessed practice quality (“How would you rate the overall quality of your meditation practice?”). Participants responded on a 1 to 10 Likert-type scale, with higher scores indicating higher practice quality. Reliability and validity of the single-item PQM were assessed in the current study.

Data Analysis

Descriptive statistics were computed first, including models examining intercorrelations between baseline variables. Multilevel models were used to account for the nesting of observations within participants, instructors, and groups (Snijders & Bosker, 2012).

Single-item PQM reliability and validity.—A series of models evaluated the validity and reliability of the single-item PQM. Test-retest reliability was assessed by examining stability across participants’ first two single-item PQM scores. Convergent validity was assessed by examining the association between single-item PQM scores provided within the first week of MBSR and the full PQM administered at baseline following the BCT. Discriminant validity was assessed by examining the association between initial single-item PQM ratings and social desirability. Construct validity was assessed by examining changes in single-item PQM scores over the course of MBSR and by examining the association between changes in single-item PQM scores and overall practice time. Predictive validity was assessed by examining the association between changes in single-item PQM scores and changes in mindfulness and psychological symptoms over the course of MBSR.

Changes in single-item PQM scores over time.—Longitudinal, multilevel models evaluated changes in single-item PQM scores over the course of MBSR using the ‘lme4’ package (Bates, Machler, Bolker, & Walker, 2015) in the R statistical software environment

(R Core Team, 2018). Models were compared using χ^2 tests to assess improvements in fit associated with the addition of a quadratic time effect and random slopes. The final model included both random intercept and random slope components (see Supplemental Materials Table 2). Random slope coefficients were extracted from the final model to reflect change in single-item PQM scores over the course of MBSR.

Multilevel mediation.—Multilevel mediation models evaluated changes in single-item PQM scores as a mediator of the link between practice time and outcome. As a formal test of mediation, estimates of the indirect effect (i.e., effect of practice time on outcome through changes in practice quality) were examined using the ‘mediate’ package with $n = 1000$ Monte Carlo draws for quasi-Bayesian approximation (Tingley et al., 2014). Standardized versions of the outcome and predictor variables (i.e., z-transformed) were used to aid in interpretation.

Sensitivity analyses.—Four sensitivity analyses assessed the robustness of the mediation findings. Models were run with unlikely practice time values retained (i.e., raw values), practice time restricted to formal practice, controlling for social desirability, and using non-parametric bootstrapped confidence intervals which relax regression assumptions (Fox & Weisberg, 2011).

Results

Descriptive Statistics

Sample descriptive statistics are reported in Supplemental Materials Table 1. Examination of correlations between study variables at baseline showed significant associations between several self-report outcomes and social desirability as well as between self-report mindfulness, practice quality, and psychological symptoms (Supplemental Materials Table 3). Behavioral mindfulness (BCT) was not associated with either self-report measure of mindfulness (FFMQ, MAAS) or with practice quality (full PQM; $r_s < .04$). Baseline BCT scores were also on average close to one (0.96), suggesting possible ceiling effects. Study variables were within recommended bounds for normality (Kim, 2013), with the exception of baseline BCT and informal practice time. Intraclass correlations representing the proportion of variance in scores occurring at the participant- (for weekly measures), group-, and instructor-levels of nesting are reported in Supplemental Materials Table 4. No evidence of group or instructor effects was found for any outcome variable ($ICCs = 1\%$). A small proportion of variance in practice time appeared at the group level, although the bulk of the variance was participant-level ($ICC = .60$ for total practice time). All non-residual variance in single-item PQM ratings was participant-level ($ICC = .55$).

Single-Item PQM Reliability and Validity

A total of 570 single-item PQM ratings were provided, with a mean of 5.94 assessment per participant ($SD = 1.43$, range = 2 to 8). The single-item PQM showed adequate test-retest reliability between the first and second administration ($r = .75$, $p < .001$). Evidence for convergent validity was found in a medium-sized positive association between the pre-treatment full PQM (administered following the BCT) and single-item PQM scores provided

within the first week of MBSR ($r = .39, p = .008$). Initial single-item PQM ratings were not associated with social desirability ($r = .11, p = .478$), supporting the measure's discriminant validity. Construct validity was supported by a moderate, positive association between changes in single-item PQM ratings and average daily practice time ($r = .48, p < .001$). Supporting the measure's predictive validity, increases in single-item PQM ratings were associated with higher post-treatment self-report mindfulness ($\beta = .35$ and $.30, ps < .001$ for FFMQ and MAAS, respectively) and lower psychological symptoms ($\beta = -.19, p = .013$) when controlling for pre-treatment. Increases in single-item PQM scores were not associated with post-treatment behavioral mindfulness ($\beta = -.02, p = .879$).

Changes in Single-Item PQM Scores Over Time

A series of nested models assessed changes in single-item PQM scores over the course of MBSR. A model including both a fixed and random effect for time provided optimum fit (superior to models without these elements as well a model with an additional quadratic term for time, see Supplemental Materials Table 5). The full final model is reported in Supplemental Materials Table 6. The model indicated a positive fixed effect for time ($B = 0.12$ per week, $p < .001$). Random slope coefficients were extracted from this model representing participant-level changes in single-item PQM scores over time.

As a longitudinal test of the relationship between single-item PQM scores and practice time, a cross-level interaction was added to the final single-item PQM longitudinal model. A significant interaction was detected between time (in weeks) and average daily practice time ($B = 0.004, t[87.64] = 2.51, p = .014$). As shown in Supplemental Materials Figure 1, greater practice time predicted a steeper increase in practice quality on the single-item PQM.

Multilevel Mediation

As noted above, a significant relationship was found between practice time and improvement in practice quality (i.e., predictor and mediator; Baron & Kenny, 1986). In addition, greater overall practice time was associated with higher post-treatment self-report mindfulness ($\beta = .28, p < .001$ and $\beta = .24, p = .003$, for FFMQ and MAAS, respectively) and lower post-treatment psychological symptoms ($\beta = -.20, p = .011$, for DASS) when controlling for pre-treatment (i.e., predictor and outcome). Practice time was not associated with post-treatment behavioral mindfulness ($\beta = .03, p = .764$, for BCT) when controlling for pre-treatment.

In multilevel mediation models, a significant indirect effect of practice time on changes in self-report mindfulness through improved practice quality was found suggesting mediation (average causal mediation effect [ACME] $\beta = .14 [.06, .24], p < .001$, and ACME $\beta = .12 [.04, .22], p < .001$, for FFMQ and MAAS, respectively). The indirect effect was not significant for psychological symptoms (ACME $\beta = -.06 [-.15, .03], p = .172$ for DASS) or behavioral mindfulness (ACME $\beta = -.02 [-.14, .10], p = .740$).

Sensitivity Analyses

Significance tests for estimates of the indirect effect of practice time through changes in practice quality were equivalent across four sensitivity analyses (Supplemental Materials

Table 7), with two exceptions: a significant indirect effect of practice time on psychological symptoms through practice quality was present for the analyses using raw practice time values or formal practice time only (AMCEs = -0.08 , $ps < .05$).

Discussion

The current study explored the possibility that improved mindfulness practice quality mediates the link between practice time and outcomes in the context of MBSR. Results corroborated two previous studies showing that practice quality predicts treatment outcome (Del Re et al., 2013; Goldberg et al., 2014). Importantly, results extended prior work showing practice quality may function as a mechanism underlying the beneficial effects of practice time, with improved practice quality acting as a mediator linking practice time and outcome. Given the novel assessment of practice quality using a single-item, reliability and validity evidence for this measure is important to consider prior to discussing the mediation results.

Several pieces of evidence supported the reliability and validity of the single-item PQM. The single-item PQM showed adequate test-retest reliability, modest convergent validity (with the full PQM at baseline), and discriminant validity (uncorrelated with social desirability). Construct validity was supported by increases in single-item PQM scores over the course of MBSR and a strong relationship between changes in single-item PQM scores and total practice time. Supporting the measure's predictive validity, increases in single-item PQM scores predicted improvements at post-treatment in several study outcomes including self-report mindfulness and psychological symptoms, although not behavioral mindfulness. The lack of association with improvements in behavioral mindfulness may reflect a limitation of the single-item PQM. It may also reflect limitations of the BCT, which was uncorrelated with self-report mindfulness at baseline in the current sample and appeared at risk for ceiling effects. However, taken together, the single-item PQM appeared to have generally desirable psychometric features.

Evidence for reliability and validity of the single-item PQM is notable particularly given single-item measures are likely to be less reliable (i.e., more confounded with error variance; Crocker & Algina, 2008) attenuating tests of validity. Reasonable concerns about the use of single-item measures have been made, and simulation studies suggest that single-item scales rarely match the predictive validity of multi-item scales (Diamantopoulos, Sarstedt, Fuchs, Wilczynski, & Kaiser, 2012). However, single-item scales with strong psychometric properties have been developed (e.g., de Boer et al., 2004) and their use has been recommended within intensive sampling paradigms (e.g., daily diary, experience sampling; Fisher & To, 2012).

Longitudinal models further clarified changes in single-item PQM scores over the course of MBSR. A significant random slope coefficient indicated that participants varied in their rate of change in single-item PQM scores over the course of MBSR. Supporting a theory-specified link between practice time and practice quality, a cross-level interaction indicated that change in single-item PQM scores were steeper for participants with higher average practice time.

Finally, multilevel mediation models indicated an indirect effect of practice time on post-treatment self-reported mindfulness mediated through improvements in practice quality. This effect was robust across four sensitivity analyses, suggesting it is not driven by treatment of practice time outliers, inclusion of informal practice, social desirability, or violation of regression assumptions. Practice quality did not mediate effects on psychological symptoms in the primary model, although an indirect effect was detected when using raw practice time values and restricting to formal practice. The possibility that certain kinds of practice (i.e., formal vs. informal) relate differently to practice quality could be explored further. There was no evidence for mediation effects on behavioral mindfulness.

Taken together, our findings suggest practice quality is an important mechanism within the context of MBIs given changes in single-item PQM scores mediated the link between practice time and two commonly used self-report measures of mindfulness (FFMQ, MAAS). Although not without substantial psychometric and conceptual limitations (Goldberg et al., in press), self-report mindfulness appears to tap psychological tendencies at least partially responsible for the therapeutic effects of MBSR and MBCT (Gu et al., 2015). A process variable that predicts improvements on these measures may be of scientific and clinical value.

Several research and clinical implications are worth noting. From a clinical standpoint, this study suggests that, similar to the development of other skills (Ericsson, 2008), the quality of one's mindfulness practice matters. Thus, rather than solely encouraging participants to engage in a certain *quantity* of practice, emphasis should also be placed on engaging in high *quality* practice. Within the context of MBSR, in which participants are often relatively naive to meditation practice, this deliberate, high quality practice may primarily involve applying appropriate effort during meditation practice (Bodhi, 2005) and gently discouraging the mind from freely engaging in habitual mental tendencies (e.g., mind wandering). Encouraging participants to attend to the quality of their practice may need to be done delicately, so as not to generate counter-productive attitudinal stances (e.g., harsh self-judgment, strained effort). At more advanced stages of practice (i.e., long-term meditation practitioners), the meaning of practice quality may shift, perhaps becoming less deliberate and more intuitive as mental habits are more deeply inculcated and an individual is better able to effortlessly monitor performance (Ericsson, 2008). Of course, the current results do not suggest that practice time is unimportant, but rather its importance may occur through its association with improved practice quality. Practice quality may be a useful clinical feature for MBSR participants and instructors to track, in keeping with routine outcome monitoring and measurement-based care (Wampold, 2015). Stagnant or deteriorating practice quality could alert instructors and participants of difficulties that need to be addressed.

From a research standpoint, our results support future work examining this aspect of MBIs. Results suggest that practice quality may be a modifiable mechanism that influences outcomes within MBIs (i.e., changes in self-report mindfulness and psychological symptoms). However, research in this area is nascent. Researchers studying MBIs could consider including measures of practice quality in future studies, particularly within clinical samples in which symptom change is more likely (e.g., MBCT for depression). The single-

item measure used here may be one minimally burdensome measure to consider, although a multi-item scale such as the full PQM (Del Re et al., 2013) may be preferred (Diamantopoulos et al., 2012).

It would also be valuable to more thoroughly interrogate the relationship between practice time, practice quality, and outcome, for example through randomly assigning participants to dosage conditions. More frequent assessment of practice quality (e.g., after each period of formal meditation) and outcomes (e.g., through experience sampling paradigms) may help further unpack these relationships. In particular, such a design could help clarify causal relationships between practice quality and outcome. Broadly, it would be useful to better understand what factors influence participants' ratings of practice quality. It may be that practice quality is partially an epiphenomenon of changes in psychological symptoms and increases in mindfulness (i.e., practice is perceived as high quality when these changes are occurring). This possibility of bidirectional or reverse causality is akin to issues raised in the therapeutic alliance literature (Fluckiger, Del Re, Wampold, & Horvath, 2018) and could be explored using intensive longitudinal data (e.g., Falkenström, Granström, & Holmqvist, 2013).

Future studies could also include additional objective outcome measures or mindfulness teacher-rated measures of practice quality to rule out the possibility that links between practice quality and outcome are driven by mono-method bias (e.g., social desirability). It would also be helpful to validate self-report practice quality in a laboratory setting with both behavioral (e.g., mind-wandering probes) and psychophysiological measures (e.g., autonomic arousal, stress hormones) drawn from the nomological network of practice quality (Levinson et al., 2014). Experience sampling paradigms could assess practice quality in daily life. Lastly, it may be worthwhile exploring the nature of practice quality among long-term practitioners, for whom practice may function differently (Ericsson, 2008).

The current study has several important limitations. First, we relied heavily on self-report measures of practice quality, practice time, and outcome. As noted, this may have inflated the observed associations through response set biases (e.g., social desirability). This potential confound may have been attenuated for changes in single-item PQM scores which were derived from random slope coefficients (presumably less susceptible to efforts to respond in a socially desirable way). Further, results were unchanged in models controlling for social desirability. While we included a behavioral measure purported to measure mindfulness to further assess this source of bias, the lack of association between the BCT and self-report mindfulness at baseline as well as potential ceiling effects on the measure raises questions regarding its construct validity. Second, as noted previously (Goldberg et al., 2014), self-assessment of practice quality may be difficult and unreliable, particularly when an individual is less familiar with mindfulness practice and may also have lower meta-awareness (Grossman, 2011). Our use of a single-item practice quality measure may have exacerbated these psychometric limitations and further reduced reliability and validity. Similarly, our single-item measure did not include key aspects of deliberate practice (e.g., opportunities for feedback), making it unclear whether it can be construed as assessing deliberate practice *per se*. Third, practice time and quality were assessed weekly (rather than daily, or per meditation practice period), which may have introduced recall biases and error

(Shiffman, Stone, & Hufford, 2008). Fourth, a portion of the sample did not complete weekly questionnaires and/or post-treatment measures and were thus excluded from some analyses. Although those who withdrew did not differ from those retained in the study on baseline demographic or outcome measures and attrition is to be expected in a primarily remote study design (e.g., Arean et al., 2016), it is still conceivable that attrition may have influenced the pattern of findings. Fifth, while larger than previous studies of practice quality, the current sample size was relatively modest and true effects may have gone undetected due to Type II error. Sixth, the sample was predominantly white, educated, and female, which may limit generalizability to other demographics. It is vital that future MBI research include more diverse participant samples. Seventh, our use of a non-clinical sample may have restricted the range of psychological symptoms and decreased statistical power. Eighth, our study lacked a follow-up assessment point, so the persistence of the observed patterns is unclear.

These limitations notwithstanding, the current study extends prior research on practice quality within MBIs by demonstrating improvements in practice quality mediate the link between practice time and outcome. Thus, practice time appears impactful in the extent to which it leads to improved practice quality. This work adds to efforts to optimize the delivery of MBIs within health care (Crane & Kuyken, 2013), settings in which questions regarding the necessary dosage of home practice may directly impact the acceptability of treatment. Further, this study introduces a single-item practice quality measure that could be monitored routinely, providing MBI instructors and participants information predictive of post-treatment outcomes.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

This work was supported by the National Center for Complementary and Alternative Medicine (NCCAM) P01AT004952 to Richard J. Davidson and Mind and Life Institute Varela Award to Simon B. Goldberg. Support for this research was also provided by the University of Wisconsin-Madison, Office of the Vice Chancellor for Research and Graduate Education with funding from the Wisconsin Alumni Research Foundation. Any views, findings, conclusions, or recommendations expressed in this publication do not necessarily reflect those of the Mind & Life Institute. Richard J. Davidson is the founder, president, and serves on the board of directors for the non-profit organization, Healthy Minds Innovations, Inc.

References

- Adams ZW, Sieverdes JC, Brunner-Jackson B, Mueller M, Chandler J, Diaz V, ... & Treiber FA (2018). Meditation smartphone application effects on prehypertensive adults' blood pressure: Dose-response feasibility trial. *Health Psychology, 37*(9), 850–860. doi: 10.1037/hea0000584 [PubMed: 30010353]
- Arean PA, Hallgren KA, Jordan JT, Gazzaley A, Atkins DC, Heagerty PJ, & Anguera JA (2016). The use and effectiveness of mobile apps for depression: Results from a fully remote clinical trial. *Journal of Medical Internet Research, 18*(12).
- Baer RA, Smith GT, Hopkins J, Krietemeyer J, & Toney L (2006). Using self-report assessment methods to explore facets of mindfulness. *Assessment, 13*, 27–45. [PubMed: 16443717]

- Baron R & Kenny D (1986). The moderator-mediator variable distinction in social psychological research: Conceptual, strategic, and statistical considerations. *Journal of Personality and Social Psychology*, 51(6), 1173–1182. [PubMed: 3806354]
- Bates D, Machler M, Bolker B, & Walker S (2015). Fitting Linear mixed-effects models using lme4. *Journal of Statistical Software*, 67(1), 1–48. doi:10.18637/jss.v067.i01
- Bodhi B (2005). *In the Buddha's words: An anthology of discourses from the Pali canon*. Somerville, MA: Wisdom Publications, Inc.
- Brown TA, Chorpita BF, Korotitsch W, & Barlow D (1997). Psychometric properties of the Depression Anxiety Stress Scales (DASS) in clinical samples. *Behaviour Research and Therapy*, 35(1), 79–89. [PubMed: 9009048]
- Brown K & Ryan R (2003). The benefits of being present: Mindfulness and its role in psychological well-being. *Journal of Personality and Social Psychology*, 84(4), 822–848. [PubMed: 12703651]
- Burke A, Lam CN, Stussman B, & Yang H (2017). Prevalence and patterns of use of mantra, mindfulness and spiritual meditation among adults in the United States. *BMC Complementary and Alternative Medicine*, 17(316). doi: 10.1186/s12906-017-1827-8
- Clarke TC, Barnes PM, Black LI, Stussman BJ, & Nahin RL (2018). *Use of yoga, meditation, and chiropractors among U.S. adults aged 18 and over*. Hyattsville, MD: National Center for Health Statistics.
- Crane RS, & Kuyken W (2013). The implementation of mindfulness-based cognitive therapy: Learning from the UK health service experience. *Mindfulness*, 4(3), 246–254. doi: 10.1007/s12671-012-0121-6 [PubMed: 23956806]
- Crawford JR, & Henry JD (2003). The Depression Anxiety Stress Scales (DASS): Normative data and latent structure in a large non-clinical sample. *British Journal of Clinical Psychology*, 42(2), 111–131. [PubMed: 12828802]
- Crocker L, & Algina J (2008). *Introduction to classical and modern test theory*. Mason, Ohio: Cengage Learning.
- de Boer AG, van Lanschot JJ, Stalmeier PF, van Sandick JW, Hulscher JB, de Haes JC, & Sprangers MA (2004). Is a single-item visual analogue scale as valid, reliable and responsive as multi-item scales in measuring quality of life? *Quality of Life Research*, 13(2), 311–320. [PubMed: 15085903]
- Del Re AC, Flckiger C, Goldberg SB, & Hoyt WT (2013). Monitoring mindfulness practice quality: An important consideration in mindfulness practice. *Psychotherapy Research*, 23(1), 54–66. doi: 10.1080/10503307.2012.729275 [PubMed: 23046287]
- Demarzo MMP, Cebolla A, & Garcia-Campayo J (2015). The implementation of mindfulness in healthcare systems: a theoretical analysis. *General Hospital Psychiatry*, 37(2), 166–171. doi: 10.1016/j.genhosppsy.2014.11.013 [PubMed: 25660344]
- Diamantopoulos A, Sarstedt M, Fuchs C, Wilczynski P, & Kaiser S (2012). Guidelines for choosing between multi-item and single-item scales for construct measurement: a predictive validity perspective. *Journal of the Academy of Marketing Science*, 40(3), 434–449. doi: 10.1007/s11747-011-0300-3
- Dimidjian S, & Segal ZV (2015). Prospects for a clinical science of mindfulness-based intervention. *American Psychologist*, 70(7), 593–620. doi: 10.1037/a0039589 [PubMed: 26436311]
- Ericsson KA (2008). Deliberate practice and acquisition of expert performance: A general overview. *Academic Emergency Medicine*, 75(11), 988–994.
- Falkenström F, Granström F, & Holmqvist R (2013). Therapeutic alliance predicts symptomatic improvement session by session. *Journal of Counseling Psychology*, 60(3), 317–328. doi: 10.1037/a0032258 [PubMed: 23506511]
- Fisher CD, & To ML (2012). Using experience sampling methodology in organizational behavior. *Journal of Organizational Behavior*, 33(7), 865–877. doi: 10.1002/job.1803
- Flckiger C, Del Re AC, Wampold BE, & Horvath AO (2018). The alliance in adult psychotherapy: A meta-analytic synthesis. *Psychotherapy*, 55(4), 316–340. doi: 10.1037/pst0000172 [PubMed: 29792475]
- Fox J & Weisberg S (2011). *An R companion to applied regression (2nd ed.)*. Thousand Oaks CA: Sage.

- Goldberg SB, Del Re AC, Hoyt WT, & Davis JM (2014). The secret ingredient in mindfulness interventions? A case for practice quality over quantity. *Journal of Counseling Psychology*, 67(3), 491–497. doi: 10.1037/cou0000032
- Goldberg SB, Tucker RP, Greene PA, Davidson RJ, Wampold BE, Kearney DJ, & Simpson TL (2018). Mindfulness-based interventions for psychiatric disorders: A systematic review and meta-analysis. *Clinical Psychology Review*, 59, 52–60. [PubMed: 29126747]
- Goldberg SB, Tucker RP, Greene PA, Simpson TL, Hoyt WT, Kearney DJ, & Davidson RJ (in press). What can we learn from randomized clinical trials about the construct validity of self-report measures of mindfulness? A meta-analysis. *Mindfulness*. doi: 10.1007/s12671-018-1032-y
- Grossman P (2011). Defining mindfulness by how poorly I think I pay attention during everyday awareness and other intractable problems for psychology's (re)invention of mindfulness: Comment on Brown et al. (2011). *Psychological Assessment*, 23(4), 1034–1040. doi: 10.1037/a0022713 [PubMed: 22122674]
- 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. doi: 10.1016/j.cpr.2015.01.006 [PubMed: 25689576]
- Kabat-Zinn J (2013). *Full catastrophe living: Using the wisdom of your body and mind to face stress, pain, and illness* (Revised ed.). New York: Random House.
- Levinson DB, Stoll EL, Kindy SD, Merry HL, & Davidson RJ (2014). A mind you can count on: Validating breath counting as a behavioral measure of mindfulness. *Frontiers in Psychology*, 5(1202), 1–10. doi:10.3389/fpsyg.2014.01202 [PubMed: 24474945]
- Lovibond SH & Lovibond PF (1995). *Manual for the Depression Anxiety Stress Scales* (2nd. ed.). Sydney, Australia: Psychology Foundation.
- Parsons CE, Crane C, Parsons LJ, Fjorback LO, & Kuyken W (2017). Home practice in Mindfulness-Based Cognitive Therapy and Mindfulness-Based Stress Reduction: A systematic review and meta-analysis of participants' mindfulness practice and its association with outcomes. *Behaviour Research and Therapy*, 95, 29–41. [PubMed: 28527330]
- R Core Team (2018). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria URL <https://www.R-project.org/>.
- Tingley D, Yamamoto T, Hirose K, Keele L, & Imai K (2014). mediation: R package for causal mediation analysis. *Journal of Statistical Software*, 59(5), 1–38. [PubMed: 26917999]
- Reynolds W (1982). Development of reliable and valid short forms of the Marlowe-Crowne Social Desirability Scale. *Journal of Clinical Psychology*, 35(1), 119–125.
- Shiftman S, Stone AA, & Hufford MR (2008). Ecological momentary assessment. *Annual Review of Clinical Psychology*, 4, 1–32. doi: 10.1146/annurev.clinpsy.3.022806.091415
- Snijders TA & Bosker RJ (2012). *Multilevel analysis: An introduction to basic and advanced multilevel modeling* (2nd ed.). London: Sage.
- Tingley D, Yamamoto T, Hirose K, Keele L, & Imai K (2014). mediation: R package for causal mediation analysis. *Journal of Statistical Software*, 59(5), 1–38. [PubMed: 26917999]
- Wampold BE (2015). Routine outcome monitoring: Coming of age - with the usual developmental challenges. *Psychotherapy*, 52(4), 458–462. doi: 10.1037/pst0000037 [PubMed: 26641376]
- Wielgosz J, Goldberg SB, Kral TRA, Dunne JD, & Davidson RJ (2019). Mindfulness meditation and psychopathology. *Annual Review of Clinical Psychology*, 15, 1–32. doi: 10.1146/annurev-clinpsy-021815-093423
- Wong KF, Massar SA, Chee MW, & Lim J (2018). Towards an Objective Measure of Mindfulness: Replicating and Extending the Features of the Breath-Counting Task. *Mindfulness*, 1–9. doi: 10.1007/s12671-017-0880-1

Public Significance Statement:

This study suggests that improvement in practice quality functions as a mechanism linking meditation practice time and outcomes within mindfulness-based stress reduction.

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript