



The validation of an active control intervention for Mindfulness Based Stress Reduction (MBSR)

Donal G. MacCoon^{a,*}, Zac E. Imel^b, Melissa A. Rosenkranz^{a,c}, Jenna G. Sheftel^a, Helen Y. Weng^a, Jude C. Sullivan^d, Katherine A. Bonus^e, Catherine M. Stoney^f, Tim V. Salomons^{a,c}, Richard J. Davidson^{a,c}, Antoine Lutz^a

^a Waisman Laboratory for Brain Imaging and Behavior, University of Wisconsin – Madison, 1500 Highland Avenue, Madison, WI 53705-2280, USA

^b Department of Psychiatry and Behavioral Sciences, School of Medicine and Public Health, University of Washington, Seattle, WA, USA

^c Department of Psychology, University of Wisconsin – Madison, Madison, WI, USA

^d UW Health Sports Medicine, Madison, WI, USA

^e UW Health Integrative Medicine, Mindfulness Program, Madison, WI, USA

^f National Institutes of Health: National Heart Lung and Blood Institute, Bethesda, MD, USA

ARTICLE INFO

Article history:

Received 26 May 2011

Received in revised form

26 October 2011

Accepted 27 October 2011

Keywords:

Mindfulness

Mindfulness Based Stress Reduction, MBSR

Active control

ABSTRACT

Most of the extant literature investigating the health effects of mindfulness interventions relies on wait-list control comparisons. The current article specifies and validates an active control condition, the Health Enhancement Program (HEP), thus providing the foundation necessary for rigorous investigations of the relative efficacy of Mindfulness Based Stress Reduction (MBSR) and for testing mindfulness as an active ingredient. 63 participants were randomized to either MBSR ($n = 31$) or HEP ($n = 32$). Compared to HEP, MBSR led to reductions in thermal pain ratings in the mindfulness- but not the HEP-related instruction condition ($\eta^2 = .18$). There were significant improvements over time for general distress ($\eta^2 = .09$), anxiety ($\eta^2 = .08$), hostility ($\eta^2 = .07$), and medical symptoms ($\eta^2 = .14$), but no effects of intervention. Practice was not related to change. HEP is an active control condition for MBSR while remaining inert to mindfulness. These claims are supported by results from a pain task. Participant-reported outcomes (PROs) replicate previous improvements to well-being in MBSR, but indicate that MBSR is no more effective than a rigorous active control in improving these indices. These results emphasize the importance of using an active control condition like HEP in studies evaluating the effectiveness of MBSR.

© 2011 Elsevier Ltd. All rights reserved.

Mindfulness based interventions, particularly Mindfulness-Based Stress Reduction (MBSR¹; Kabat-Zinn, 1990) are increasingly popular. There is substantial evidence that MBSR improves mental and physical health compared to wait-list controls and treatment as usual, and is of comparable efficacy to other psychological interventions (e.g., Barnhofer et al., 2007; Davidson et al., 2003; Gregg, Callaghan, Hayes, & Glenn-Lawson, 2007; Kabat-Zinn et al., 1998; Ma & Teasdale, 2004; Pradhan et al., 2007; Specia, Carlson, Goodey, & Angen, 2000). However, a complete understanding of the mechanisms by which MBSR is efficacious for these outcomes and a valid test of mindfulness as the presumed active ingredient is not currently possible due to the lack of a suitable control intervention. The validation of such a control is the subject of this article.

A direct test of the efficacy of MBSR's active ingredients requires a comparison of MBSR to an active control that matches MBSR in non-specific factors (e.g., structure) but does not contain mindfulness as an active ingredient (Grunbaum, 1986; Kirsch, 2005). There are only two studies involving MBSR-like interventions that use control conditions that approach this standard (Grossman, Tiefenthaler-Gilmer, Raysz, & Kesper, 2007; McMillan, Robertson, Brock, & Chorlton, 2002)². McMillan and colleagues randomly assigned 145 people with traumatic brain injury either to "Attention Control Training" (based on Kabat-Zinn's work but not MBSR), physical exercise, or a wait-list control and found no differences between the two active groups. Limited descriptions of interventions and providers make it difficult to evaluate if the control was adequate.

* Corresponding author.

E-mail address: dgmaccoon@wisc.edu (D.G. MacCoon).

¹ Abbreviations: MBSR = Mindfulness Based Stress Reduction; HEP = Health Enhancement Program.

² A recent study by Raison and colleagues (Pace et al., 2008) also uses an active control condition but focuses on compassion meditation rather than the mindfulness meditation taught in MBSR.

Grossman and colleagues assigned participants with fibromyalgia to MBSR ($n = 43$) or social support/relaxation ($n = 15$). MBSR participants improved relative to the control group on measures of anxiety, depression, quality of life, and pain regulation. However, the study was quasi experimental and the control condition was subject to several limitations that are common in studies evaluating specific components of behavioral interventions (Wampold et al., 1997). Specifically, patients received less contact with providers in the control condition than in the MBSR condition. In addition, the control conditions appeared to be defined more by proscriptions (e.g., “emphasis was placed upon not describing or training mindfulness skills to the control group”), rather than the skillful provision of common therapeutic elements, which may bias tests of intervention effects (Mohr et al., 2009).

An appropriate test of mindfulness as an active ingredient requires a control condition that attends to three major limitations typical of active controls in behavioral intervention research. First, since researcher allegiance to intervention is a strong predictor of differences between two interventions that are directly compared, accounting for up to 10% of the variability in treatment outcomes (Gaffan, Tsaoasis, & Kemp-Wheeler, 1995; Luborsky, Diguer, Luborsky, & Schmidt, 1999; Wampold, 2001) and up to 69% of the differences between interventions (Imel, Wampold, Miller, & Fleming, 2008; Luborsky et al., 1999), researchers have recommended balancing allegiance when two psychological interventions are directly compared (Hollon, 1999). Second, active and control interventions should be structurally equivalent. Structural variables include number and duration of sessions, therapist training and qualifications, format of the therapy (e.g., group or individual), and the ability of participants to discuss their particular problems. If interventions are unequal in these ways, differences between interventions may be a result of structural non-equivalencies rather than the mechanism of interest. Indeed, when structural differences between interventions and active controls are eliminated, differential efficacy may disappear. In a meta-analysis of 21 psychotherapy studies, the effect of treatment was Cohen’s $d = .47$ when the control was not equivalently structured and only $d = .15$ when it was (Baskin, Tierney, Minami, & Wampold, 2003). Finally, the active control should include all non-specific factors present in MBSR. Many active controls that are designed to control for non-specific factors do not contain an accepted rationale or corresponding specific ingredients and would not plausibly be offered as efficacious by providers (Wampold et al., 2010). A well-designed control should include: (a) positive expectation for intervention success by both the therapist and client (Mohr et al., 2009), (b) a therapeutic relationship, (c) provision of a plausible alternative and adaptive explanation for distress (i.e., therapeutic rationale), and (d) some corresponding action for its alleviation (i.e., specific ingredients; Frank & Frank, 1993).

The objective of the current study was to isolate mindfulness as a specific ingredient by designing a control condition that meets the criteria above, while not containing any mindfulness training. The Health Enhancement Program (HEP; MacCoon et al., 2011) was

designed to accomplish these goals. Instructors were chosen for their expertise in, and allegiance to, the class content and the mechanisms associated with its efficacy: MBSR instructors were experts in mindfulness and HEP instructors were experts in their areas (see Supplementary materials). Our laboratory’s interest in mindfulness is well-known. To help reduce the potential impact of this researcher allegiance, (a) researchers were not part of teaching the classes, (b) instructors played a major role in the design and implementation of their intervention (as previously discussed), and (c) one member of the design team (Z.I.), who played an important role in consultation regarding the rigor of HEP as an active control condition, has primary allegiance to common factor approaches to therapy and little allegiance to mindfulness (for a more detailed discussion, see Supplementary materials).

Both HEP and MBSR were structurally equivalent, having a group format, meeting once a week for 2.5 h (3 h for first and last sessions) for 8 weeks with an “all day” component (9 a.m.–4 p.m.) after week 6, and completing the same amount of home practice (45 min, 6 of 7 days each week).

HEP content met the following criteria: (1) class activities were chosen to match MBSR activities as closely as possible (see Table 1), (2) these activities represented valid, active, therapeutic ingredients in their own right, and (3) these ingredients did not include mindfulness. Thus, the purpose of walking in MBSR is to cultivate awareness in movement, whereas the purpose of walking in HEP is the cardiovascular benefits of the physical activity for cardiovascular training and followed recommendations from the Centers for Disease Control regarding intensity and frequency of physical activity (Haskell et al., 2007). Similarly, the purpose of yoga in MBSR is largely to cultivate nonjudgmental awareness of physical sensations and respecting one’s own physical limits as they change over time. In contrast, the purpose of the balance, posture, and agility exercises in HEP’s functional movement is to augment one’s physical strength, balance, agility and joint mobility resulting in a physically more resilient individual less prone to sustain injury from spontaneous or unpredictable events (e.g., tripping on a curb, slipping on icy ground, lifting a heavy object; e.g., Hu & Woollacott, 1994; McGuine & Keene, 2006). The music therapy component included an exercise that matched the body scan in several ways with a primary difference being the importance of the music as the change agent rather than MBSR’s emphasis on awareness of one’s own internal states. The nutrition component included didactic material and reading, both modalities used in MBSR but the content was not related to mindfulness.

The rationale for MBSR and HEP reflect these different active ingredients. The following is a summary of the rationale underlying MBSR: Meditative awareness is fundamental to working with problems we may have because recognizing habit patterns of mind, their impact on situations and on the body, and learning to ‘respond’ rather than simply falling into habit patterns is essential in learning skillful means of recognizing ‘problems’ and being open

Table 1
Intervention content comparison.

MBSR		HEP	
In-Class	Homework	In-Class	Homework
Body Scan	Body Scan and light reading	Music Therapy: Relax, listen to music, imagery, and drawing	Relax, listen to music, imagery, and drawing
Sitting Meditation	Body Scan, Sitting Meditation, and light reading	Nutrition Education around Food Guide Pyramid	Planning meals, tracking diet, food labels, journaling
Yoga	Alternate Yoga and Body Scan, and Sitting Meditation	Functional Movement (posture, balance, core movement)	Posture, balance, coordinated movement
Walking Meditation	Walking and other practices	Physical Activity (walk/jog, stretch)	Walking and stretching
All Day (7 h): Work with all practices, Group discussion & exercises	–	“Spa Day” (7 h): Work with all practices, Group discussion & exercises	–

to more healthy options. Scientific evidence has found that mindfulness is helpful for improving various aspects of well-being, including depression, anxiety, and sleep quality.

The following is a summary of the rationale underlying HEP: In four areas, we will help you develop new habits and reinforce new ones that are known to increase well-being: (1) Physical activity enhances one's sense of well-being, increases energy, and reduces health risks, including coronary heart disease, stroke, colon cancer and diabetes; (2) Functional movement improves posture/core strength, balance, agility and joint mobility, resulting in a physically more resilient individual less prone to sustain injury from spontaneous or unpredictable events; (3) Supportive Music and Imagery and other elements of music therapy generate positive emotions to facilitate performance on concrete tasks and is used in a group setting to create a common experience, thereby increasing relaxation-related melatonin levels (Kumar et al., 1999), enhancing immune response (Bittman et al., 2001; Wachi et al., 2007), favorably changing stress-related gene expression (Bittman et al., 2005), increasing positive mood and reducing burn out (Bittman et al., 2004); (4) Incorporating evidence-based nutrition into one's eating lifestyle will help reduce the risk of cardiovascular disease, hypertension, dyslipidemia, type 2 diabetes, overweight and obesity, osteoporosis, constipation, diverticular disease, iron deficiency anemia, oral disease, malnutrition, and some cancers (U.S. Department of Health and Human Services and U.S. Department of Agriculture, 2005).

Thus, while HEP includes specific active ingredients meant to enhance health and well-being and thus represents an active intervention, it is also suitable as an active control for MBSR because it was matched to MBSR on non-specific elements but designed without mindfulness as one of those specific ingredients.

An additional limitation of many tests of mindfulness-based interventions is the reliance on self-report questionnaires to confirm the presence of mindfulness itself as an ingredient in training (e.g., Cohen-Katz et al., 2005). Due to the demand characteristics inherent in this approach, the relative transparency of the items on such measures, and the often-present requirement to judge internal mental processes (Haefel & Howard, 2010; Nisbett & Wilson, 1977), we used a thermal pain task to test that mindfulness was present in MBSR but not HEP. Relative to more traditional self-report mindfulness questionnaires, this task reduces the requirement to judge internal processes, equalizes to a greater degree demand characteristics across both HEP and MBSR interventions, and has instructions equally transparent to both interventions.

To validate HEP as a suitable active control for MBSR, we tested the following primary hypotheses: (1) that pain reactivity would be moderated by mindfulness but not control-related instructions for participants of MBSR but not participants of HEP; (2) that MBSR participants would show decreased pain ratings over time relative to HEP participants in the relevant instruction condition, a prediction based on evidence for the analgesic effects of mindfulness (e.g., Brown & Jones, 2010; Grossman et al., 2007; Kabat-Zinn, 1982; Perlman, Salomons, Davidson, & Lutz, 2010); (3) that both interventions would show reduced participant-reported mental and physical health symptoms over time, with MBSR showing greater reductions than HEP; and (4) that these predicted effects would be moderated by home practice. Similar but exploratory predictions existed for measures ranging from stress to well-being and correlations between pain data and primary self-report and practice variables (see [Supplementary materials](#)).

Methods

Participants and procedures

Participants provided their written informed consent for study procedures that were approved by the UW-Madison Health

Sciences Internal Review Board. Participants were recruited for a study on "health and well-being" through advertisements in Madison, WI area newspapers. Advertisements offered \$475 plus a free "8-week Health Enhancement Program" or "8-week Mindfulness Based Stress Reduction Class". People were informed about study requirements and screened for exclusion/inclusion criteria (see [Table 2](#)) through telephone interviews.

After telephone screening, 94 people attended one of four information sessions in which the study was described by project scientists, the classes were described by instructors, written consent was obtained, and lab visits were scheduled. Participants were organized into two cohorts based on schedules and class size restrictions, and members of each cohort were randomized to intervention by a logistical staff member through a random-number generator at the time of assignment, and underwent identical procedures separated by approximately 4 weeks. Participants were masked to research questions and researchers were masked to intervention assignment throughout data collection.

Participants completed laboratory visits at the Waisman Laboratory for Brain Imaging and Behavior at UW-Madison within the

Table 2
Inclusion/Exclusion criteria.

Inclusion criteria

Able to lie still in the scanner for 90 min
Meets MRI safety standards
Weighs under 300 pounds
18–65 years old
Right-handed
No previous experience with meditation.
No daily practice with other mind-body techniques (e.g., yoga, tai-chi, but previous exposure to yoga is okay)
In good general health as determined by the investigator
Able to walk
Able to understand and speak English
Able to provide written consent prior to admission
Able to see without glasses (as if looking through binoculars)

Exclusion criteria

Diabetes
Peripheral vascular disease, peripheral arterial disease, Raynaud's disease, or any other diagnosed circulatory disorders
Body mass index (BMI) below 18.5
Any involuntary motor disorders
Allergic to adhesive tape
A history of problems of any kind during blood draws or needle phobia
2 or more of the following: Diagnosed hypertension, Hyperlipidemia, High cholesterol, Obesity, Smoke cigarettes, Family history of coronary or atherosclerotic disease (parents/siblings prior to age 55)
Current medical disorders that might make interpretation of scan data difficult
Has a problem with alcohol or non prescription drugs
Currently uses or plans to start medications that affect CNS function, including psychotropics, opiate medication or corticosteroids, during the last 3 months (including medications for anxiety, depression, or other psychological problems)
Takes inhaled steroids for asthma? (e.g., Fluticasone)
Takes any corticosteroids
Night shift workers (11 p.m.–7 a.m.) d/t potential disruption of cortisol level variability
Diabetes requiring insulin treatment
Has TMJ (Temporal Mandibular Joint) disorder or other problems with biting/chewing
Previous training in meditation
Currently meditates on a regular basis
Daily yoga, tai-chi, or Qigong practice
Engagement in moderate sport and recreational activities more than 5 times a week
Engagement in vigorous sport and recreational activities more than 4 times a week
Not able to attend an informational session, all class meetings, and all clinic visits

four weeks prior to the class beginning (T1), within 4 weeks after class ending (T2), and approximately 4 months following class ending (T3). Participants' home practice was tracked while the class met (between T1 and T2) and during the four months between T2 and T3. After completion of T1 measures, sixty-three participants were randomized to HEP ($n = 32$) or MBSR ($n = 31$; see Fig. 1; see Table 3 for demographic information).

Thermal pain

Pain stimuli were generated by a TSA-2001 thermal stimulator (Medoc Advanced Medical Systems, Haifa, Israel) with a 30 mm × 30 mm flat thermode applied to the inside of the left wrist. A calibration procedure identical to Salomons, Johnstone, Backonja, and Davidson (2004) was used to establish a participant's pain threshold. Temperatures ranged from 45 °C to 49 °C. There were no intervention differences in temperature used, $t(36) = -.9$, $p = .34$, $\eta^2 = .03$. After calibration, participants experienced 32 trials of thermal stimulation, divided into 8 runs of 4 trials each, with a resting period and comfort check in between. The experimental procedure for each trial is depicted in Fig. 2a. On each trial of this mixed design, participants were presented with a cue to either "notice their emotions, sensations and thoughts" (MBSR-relevant condition) or to "notice the music" (a HEP-relevant condition: music-based training was a key part of HEP). Importantly, each instruction was understandable to all participants prior to training but was also designed to prime class-specific content after training. Both order of instructions (music or sensation focus) and pain condition (hot or warm) were counterbalanced across runs. At the end of each trial, participants were presented with two 11-point Likert scales, the first measuring intensity (i.e., how hot the stimulus was, 0 = "no pain", 10 = "most intense pain tolerable"), and the second measuring unpleasantness (i.e., how much the pain bothered them, 0 = "not at all unpleasant", 10 = "extremely unpleasant").

Participant-report outcomes (PROs)

The 90-item Symptom Checklist-90-R (SCL-90-R; Derogatis, 1983) consists of nine subscales and three global scales. The Global Severity Index (GSI) provides a measure of overall psychological distress and has demonstrated sensitivity to change and adequate internal consistency (Thompson, 1989). The depression, anxiety, and hostility subscales also were used (Cronbach's $\alpha = .90$, .85, and .84 respectively; test-retest reliabilities are $r = .82$, .80, and .78 respectively).

The Medical Symptoms Checklist (MSC; Travis, 1977) measures the number of medical symptoms participants' experienced as problems in the last month. While the MSC has demonstrated sensitivity to change in past studies of MBSR (Kabat-Zinn, 1982), no further psychometric data is available.

Participants also recorded minutes and sessions of home practice, both between T1 and T2 (class practice) and between the class end and T3 (four-month practice). The former was used for tests of change from T1 to T2, while total practice (class practice plus four-month practice) was used for change from T1 to T3. Participants' expectations and experience of their class were assessed using the Experience Check Questionnaire (ECQ) with a 7-point Likert scale.

Results

Analyses for all PROs except thermal pain ratings are based on participants with complete data for the time points included in the analyses. Intent-to-treat (ITT) analyses with multiple imputation (using 5 imputed datasets) were conducted and no meaningful

differences were found between the pooled results under multiple imputation and our original results. Therefore, we do not report ITT analyses.

A univariate General Linear Model (GLM) with intervention as the between-participant variable and a point value quantifying stressful life events (Stress Points) at T1 as the dependent measure indicated that interventions did not differ in terms of stressful life events, $F(1, 55) = 1.73$, $\eta^2 = .03^3$.

Outlier participants were identified based on extreme data (>3 interquartile ranges from the mean in one time point and >2 interquartile ranges from the mean on at least one other time point). Analyses were conducted with and without these outliers. Results were similar, but divergent results are highlighted when they occur. Because multiple regression is particularly susceptible to "high influence points" (e.g., Stevens, 1984), practice data identified as having a Cook's distance ≥ 1 (Cook & Weisberg, 1982) were excluded from primary analyses.

HEP and MBSR equivalence: non-specific factors

Participants entered the study with a "somewhat strong" preference ($M = 3.87$, where 4 is "somewhat strong" on the 7-point Likert scale) to be randomized to MBSR (64% preferred) over HEP (15% preferred; 21% had no preference). Furthermore, a multivariate test performed on 10 other ratings of the intervention's value revealed significant intervention differences with higher ratings for MBSR, $F(1, 53) = 3.79$, $p = .001$, $\eta^2 = .44$.

HEP and MBSR equivalence: structural equivalence

HEP and MBSR were structurally equivalent. A univariate GLM with intervention and cohort as the between-participant variables revealed no effect of intervention on drop status, $F(1, 53) = .01$, $\eta^2 < .001$, or cohort, $F(1, 53) = .01$, $\eta^2 < .001$. A similar analysis revealed no main effect of intervention, $F(1, 53) = 1.22$, $\eta^2 = .02$, or cohort, $F(1, 53) = 3.67$, $p = .06$, $\eta^2 = .07$ on number of classes attended or time spent in class, $F(1, 53) = .19$, $\eta^2 = .004$ and $F(1, 53) = 1.61$, $\eta^2 = .03$ for intervention and cohort respectively.

Time spent in formal home practice also was tested. Three participants indicated only informal practice. One of these participants was an extreme outlier on all practice metrics at all time points (e.g., reporting practice of 4–5 h per day during class). As this practice is likely misreported, this individual was not included in any analyses involving practice. The other two participants were not included in primary analyses⁴.

Univariate GLMs with intervention and cohort as between-participant variables and class practice minutes as the dependent variable revealed no significant effect of intervention, $F(1, 52) = .17$, $\eta^2 = .003$, or cohort, $F(1, 52) = 1.22$, $\eta^2 = .02$. On average, 1849 min of homework were completed during an 8-week class (about 44 min of for 6 of 7 days per week compared to 45 min assigned). Similar analyses using number of practice sessions as the dependent measure revealed a significant effect of intervention, $F(1,$

³ One analytic strategy would be to use Stress Points at T1 as a covariate. However, since no MBSR study to date has followed this strategy and we wish our results to be understood in the context of the extant literature, we opted not to do so.

⁴ Each of these participants were from the HEP intervention. Analyses were conducted with these individuals' estimated formal practice. Missing formal data were imputed by calculating the percent of formal practice out of informal practice for each participant in the study. Missing information for each of the two participants was calculated by multiplying their informal practice data by their group's mean percent formal practice for minutes and number of sessions. Results with these estimated data did not differ from primary analyses.

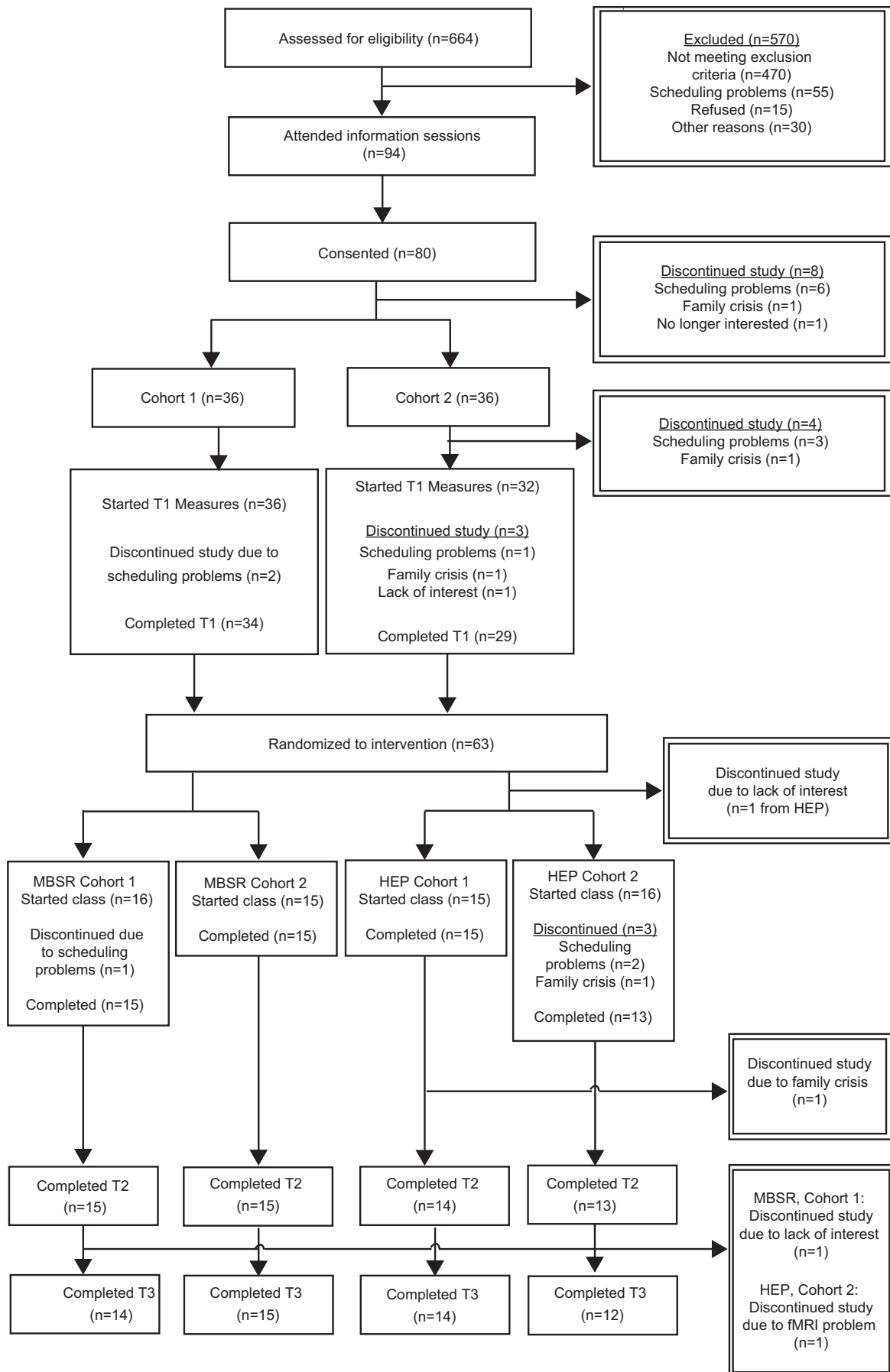


Fig. 1. CONSORT diagram detailing retention rates by study phase and reasons for dropouts.

Table 3
Demographics by intervention.

Intervention	Gender	Age		Ethnicity			
	<i>n</i> (% female)	<i>M</i> (<i>SD</i>)	Range	Caucasian	AA	Asian	Indian
HEP (<i>n</i> = 27)	23 (85.2%)	47.5 (9.3)	19–59	23 (85.2%)	3 (11.1%)	1 (3.7%)	0 (.0%)
MBSR (<i>n</i> = 30)	24 (80.0%)	44.5 (12.2)	21–59	24 (80.0%)	1 (3.3%)	2 (6.5%)	2 (6.7%)

AA = African American.

52) = 21.55, $p < .001$, $\eta^2 = .29$ ($M = 95.60$ sessions for HEP and $M = 61.19$ sessions for MBSR). There was no effect of cohort, $F(1, 52) = 2.14$, $\eta^2 = .04$.

The same analyses were conducted for total minutes of practice and revealed no effect of intervention, $F(1, 52) = 2.41$, $\eta^2 = .04$, or cohort, $F(1, 52) = 1.09$, $\eta^2 = .02$. On average, participants completed 4394 min of total practice, corresponding to about 25 min of daily practice for 6 out of 7 days per week of each month during their four-month practice (between T2 and T3). There was a significant effect of intervention for number of practice sessions, $F(1, 52) = 13.03$, $p = .001$, $\eta^2 = .20$, indicating that HEP participants completed more practice sessions than MBSR participants ($M = 225.55$ and 137.60 sessions respectively). There was no effect of cohort, $F(1, 52) = 2.56$, $p = .12$, $\eta^2 = .05$. Due to the lack of cohort effects in the above analyses, cohort is not included as a factor in subsequent analyses.

Efficacy of current MBSR

Did studying MBSR change its efficacy? The 2.8% (8 of 286⁵) drop-out rate recorded for historical UW Health MBSR data is similar to the current study's drop-out rate of 3.2% (1 person). To test the effectiveness of the current study's MBSR classes, we conducted a repeated-measures GLM on MBSR participants only with T1 and T2 GSI scores as a repeated, within-participant variable. This analysis revealed a main effect of time, $F(1, 28) = 4.62$, $p = .04$, $\eta^2 = .14$, indicating significant improvement on the GSI over time ($M = .35$ and $M = .23$ for T1 and T2 respectively). A similar analysis conducted with T1 and T2 MSC scores also revealed a main effect of time, $F(1, 29) = 7.20$, $p = .01$, $\eta^2 = .20$ ⁶.

We next compared the outcomes from the current study to those normally achieved by the UW Health MBSR classes by restricting the range of T1 GSI scores (and separately MSC scores) in the historical database to that of the current study's MBSR participant T1 GSI and MSC scores. We conducted a Monte Carlo study on the resulting historical sample ($N = 606$ for GSI and 611 for MSC) by taking random samples of 29 participants (for GSI; 30 participants for MSC) from the historical database, calculating the mean and standard deviation of the T1 outcome measure of interest (either GSI or MSC) and selecting the first 200 samples with comparable means and standard deviations for each T1 outcome measure from the current study. We then rank ordered the mean difference between T1 and T2 outcome for each of these 200 samples and compared the same mean difference for the current study to this distribution. The current study ranks of 191 (GSI) and 134 (MSC) fell within the middle 95% range of the historical distribution, indicating no significant difference between the current study's MBSR class efficacy and historical efficacy from the same institution.

⁵ Total *N* is substantially lower than the full database because only with recent data were drop-outs assessed directly.

⁶ Since the goal of these analyses is to compare our results to those in the literature, analyses with all participants included is appropriate. When GSI outliers were removed, the main effect of time was weakened to a trend, $F(1, 26) = 3.72$, $p = .07$, $\eta^2 = .13$. Results did not differ when MSC outliers were removed.

Thermal pain ratings

Out of an initial group of 43 participants with data at T2, four participants (4 MBSR) were excluded from analyses based on a priori criteria: one participant was a low rating outlier at 3 time points, one had a higher intensity rating during warm than the hot condition, and two showed a bias toward the HEP-related instruction at T1⁷. Thus, 39 participants were available for analyses across T1 and T2 (21 MBSR, 18 HEP). Because data were not available for three other participants at T3 (2 MBSR, 1 HEP), analyses across three time points involved 36 participants (20 MBSR, 16 HEP). We conducted a repeated-measures GLM with intervention as a between-participant variable and T1, T2, and T3 as a repeated, within-participant variable⁸. The dependent variable was averaged intensity and unpleasantness pain ratings in response to hot stimuli for the HEP-relevant instruction condition subtracted from the MBSR-relevant condition⁹. A significant intervention \times time interaction, $F(2, 33) = 3.6$, $p = .04$, $\eta^2 = .18$, indicates that the mindfulness (but not HEP-relevant) condition moderated pain ratings for MBSR participants relative to HEP participants (see Fig. 2b). A similar analysis for T1 and T2 replicated this finding, $F(1, 37) = 6.17$, $p = .02$, $\eta^2 = .14$. Analyses for simple effects showed significant change over time (T1, T2, and T3) for the MBSR group, $F(2, 18) = 8.5$, $p = .002$, but not the HEP group, $F(2, 16) < 1.0$. For the MBSR group, the mindfulness (but not HEP-relevant) condition decreased pain ratings at T2 compared to T1, paired *t*-test, $t(1, 20) = 4.3$, $p < .001$, and decreased at T3 compared to T1, paired *t*-test, $t(1, 19) = 2.8$, $p = .01$. The two interventions did not differ at T1, *t*-test, $t(1, 39) = 1.1$, $p = .30$, but differed at T2, $t(1, 39) = -2.4$, $p = .03$, and at T3, $t(1, 36) = -2.2$, $p = .03$.

⁷ Results remain unchanged when these two participants were included in the analysis.

⁸ We considered a variety of ways to implement mixed effects/multilevel models for our data analyses instead of repeated-measures GLM. However, despite the potential benefits of these models for characterizing effects in terms of growth parameters, and accounting for cohort effects, no sensible implementation of these models yielded a better description of our data than the approach reported. Specifically, the typical advantages of a trajectory-based HLM model (with random slopes and intercepts) are difficult for us to achieve due to the use of three time points, combined with the expected and observed nonlinearity in change over time. This is true when using (1) a random intercept only model which unrealistically constrains change over time to be the same across participants, (2) a linear random slope and intercept model, which is mis-specified because we expect and observe nonlinear trends across our three time points, (3) various nonlinear models with random intercept and slope to account for the nonlinear change out to the third time point, but with some aspect of change fixed to avoid exhausting degrees of freedom (this last feature is needed to make a level-1 residual variance estimable, as a quadratic growth curve with three random components will perfectly fit all three data points) were not realistic or did not better describe the data than the GLM analyses.

⁹ Correlations between pain intensity and unpleasantness were comparable in the MBSR and HEP participants and with *r*-values ranging between .85 and .98. We thus averaged across pain intensity and unpleasantness. Interestingly, other evidence suggests that among very long-term meditation practitioners, differences can be observed between pain intensity and unpleasantness (see Perlman et al., 2010).

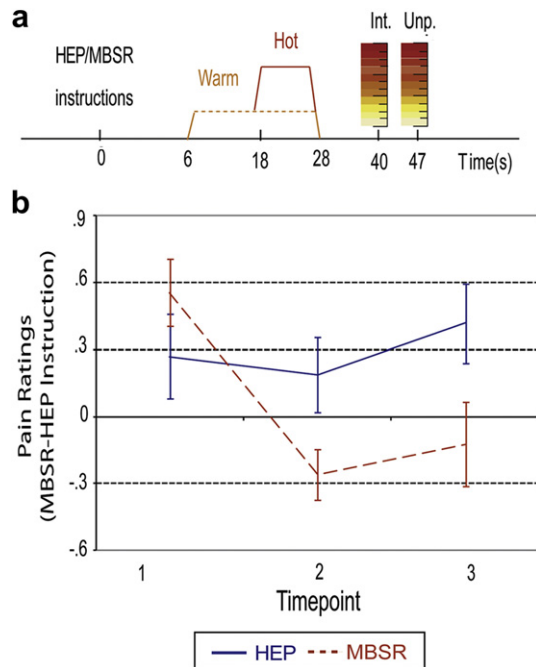


Fig. 2. (a) Method details for pain task trials, including Likert scales rating Intensity (Int) and Unpleasantness (Unp) of thermal stimuli. (b) Intervention \times Time interaction for averaged pain intensity and pain unpleasantness ratings of hot stimuli for the Control instruction condition subtracted from the Mindfulness instruction condition.

SCL-90-R

To test the effects of intervention, time, and their interaction, repeated-measures GLMs were calculated on the GSI using intervention as a between-participant variable and T1, T2, and T3 as a repeated, within-participant variable¹⁰. A significant main effect of time, $F(1, 48) = 4.90$, $p = .01$, $\eta^2 = .09$, indicated that GSI decreased. There was also a significant time \times intervention interaction, $F(1, 48) = 3.74$, $p = .04$, $\eta^2 = .07$ ($M = .28, .22, .14$ for T1, T2, and T3 respectively for HEP; $M = .26, .17, .24$ for T1, T2, and T3 respectively for MBSR; see Fig. 3a). Specific contrasts indicated a significant time \times intervention interaction between T2 and T3, $F(1, 48) = 10.51$, $p = .002$, $\eta^2 = .18$ ¹¹, and no effect between T1 and T2, $F(1, 48) = .34$, $\eta^2 = .007$. Analyses of simple effects showed no significant group differences at any time point (all F s < 1.04 , all $\eta^2 < .02$).

Analogous analyses with depressed symptoms revealed a significant time \times intervention interaction, $F(1, 48) = 4.72$, $p = .01$, $\eta^2 = .09$ ($M = .33, .34, .16$ for T1, T2, and T3 respectively for HEP; $M = .35, .21, .33$ for T1, T2, and T3 respectively for MBSR; see Fig. 3b). Specific contrasts revealed a significant intervention \times time interaction between T2 and T3, $F(1, 48) = 10.69$,

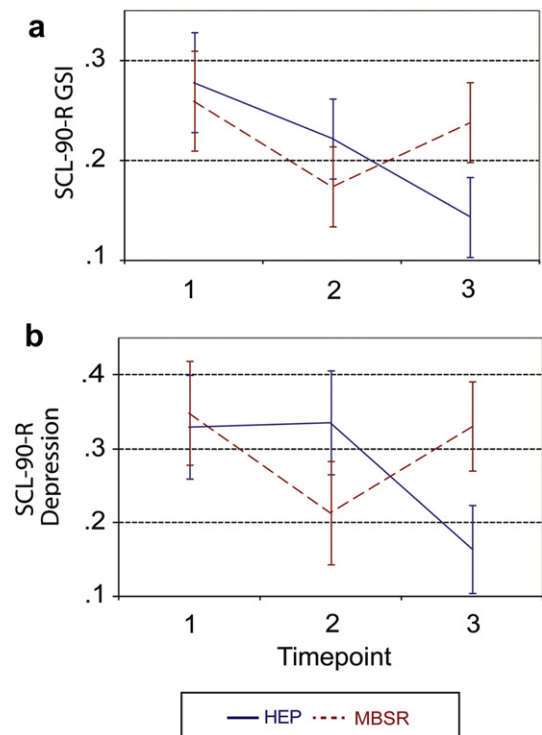


Fig. 3. (a) Intervention \times Time interaction for the SCL-90-R GSI scale. Error bars represent one standard error above and below the mean. (b) Intervention \times Time interaction for the SCL-90-R Depression sub-scale.

$p = .002$, $\eta^2 = .18$, indicating HEP participants showed decreasing symptoms of depression from T2 to T3 relative to MBSR participants who showed increasing depressive symptoms over the same time period. There was no intervention \times time interaction between T1 and T2, $F(1, 48) = 2.89$, $p = .10$, $\eta^2 = .06$. Simple effects indicated no group differences at any time point (T1 and T2 F s < 1 , all $\eta^2 < .009$; T3, $F(1, 48) = 2.16$, $\eta^2 = .04$).

With anxious symptoms as the dependent measure, analyses revealed only a significant main effect of time, $F(1, 48) = 4.19$, $p = .02$, $\eta^2 = .08$, indicating that anxious symptoms decreased over time. With symptoms of hostility as the dependent measure, there was only a significant main effect of time, $F(1, 48) = 3.58$, $p = .04$, $\eta^2 = .07$, indicating that hostility scores decreased over time¹².

Medical Symptom Checklist (MSC)

A similar repeated-measures GLM using the MSC revealed only a significant main effect of time, $F(1, 50) = 8.00$, $p = .001$, $\eta^2 = .14$, indicating decreased medical symptoms over time with improvement occurring between T1 and T2¹³.

Effects of practice

Hierarchical linear regressions assessed the impact of practice on SCL-90-R and MSC scale changes over time. Practice variables

¹⁰ In all GLM analyses involving within-participant factors of more than two levels, we report Huynh-Feldt-corrected p -values and uncorrected degrees of freedom to address violations of sphericity assumptions. Intraclass correlations (ICC) for group dependence [four groups defined by the combination of intervention and cohort (2 per intervention)] were not significant and were effectively zero for most of our outcomes. However, for the GSI and Hostility scales of the SCL-90R, the ICCs were .230 and .396 respectively. While not statistically significant, these estimated ICCs are of decent size (indeed, significant ICCs have been found using a larger dataset of historical MBSR data, see Imel, Baldwin, Bonus, & MacCoon, 2008); unfortunately, we are naturally underpowered to test intervention effects at the cohort level. Thus, we present our results with caution noting that cohort effects may exist for these outcomes.

¹¹ All effects were similar when outliers were included, except that there was no omnibus time \times intervention interaction, $F(1, 51) = 1.96$, $\eta^2 = .04$.

¹² Effects were similar when outliers were included, except that the main effect of time was weakened to a trend, $F(1, 51) = 2.68$, $p = .08$, $\eta^2 = .05$. Pain ratings were not associated with changes in self-report either from T1 to T2 (highest $r = -.27$) or from T1 to T3, (highest $r = -.28$). These results did not differ when imputed subjects were not included, except that the association between pain ratings and the MSC from T1 to T3 became a trend-level effect ($r = -.39$, $p = .06$).

¹³ Three extreme outliers were removed from analyses with similar results when included.

were mean centered, as was any T1 self-report measure. Either the T2 or T3 self-report variable of interest was the criterion, with the relevant T1 self-report measure entered in the first step, practice entered in the next step, intervention entered in the third step, and the intervention \times practice interaction entered in the fourth step. These analyses revealed no significant main effects of practice and no significant intervention \times practice interactions for either of the practice metrics for any measure from T1 to T2 or from T1 to T3 (R^2 's $\leq .06$).

Discussion

This is the first study comparing MBSR to an active control condition that was designed to be inert with respect to mindfulness, while being structurally equivalent to MBSR and credible to both patients and providers. The fact that an MBSR-relevant instruction condition moderated pain ratings relative to HEP-relevant instructions in the MBSR participants compared to the HEP participants (see Fig. 2) suggests that mindfulness was, indeed, an active ingredient in MBSR but not in HEP (hypothesis 1)¹⁴. Furthermore, consistent with extant data (e.g., Brown & Jones, 2010; Perlman et al., 2010), the same result indicates that MBSR selectively alters the unpleasantness of painful stimuli relative to HEP in the relevant instruction condition suggesting an analgesic effect of MBSR (hypothesis 2). Specifically, MBSR participants' pain ratings decrease over time whereas HEP participants' pain ratings do not change. Thus, following a mindfulness-related instruction is more effective in reducing pain than following HEP-related instructions given the same amount of exposure and training to those respective practices. This result suggests that a mindfulness-based practice may be superior for regulating pain than an approach based on music and fitness. Analyses of self-report mental and medical symptoms suggest that HEP and MBSR were effective in reducing symptoms over time, but provided little evidence of differential efficacy of one intervention over the other (contrary to hypothesis 3). There were no significant group effects for any primary outcome measure on the SCL-90-R. Furthermore, significant group \times time interactions suggest that HEP may have been superior for some outcomes. Specifically, HEP participants showed decreasing mental distress (GSI) from T2 to T3 whereas MBSR participants showed increasing mental distress over the same time period (see Fig. 3a). This result should be treated with caution, however, because of a .230 intraclass correlation (ICC) for cohort. Though underpowered for cohort-level effects, this relatively large ICC indicates that symptom reduction may depend on cohort. A similar effect was also evident for symptoms of depression (see Fig. 3b).

Contrary to our fourth hypothesis, there were no significant main effects of practice nor any group \times practice interactions for any measure from the SCL-90-R or medical symptoms (MSC), a finding consistent with past research (e.g., Davidson et al., 2003, but see Speca et al., 2000).

In addition to the lack of group differences reported for hypothesis 3 and 4, there are other indications that HEP and MBSR were equivalent. Both interventions were rated favorably by participants, had similar drop-out rates, attendance, homework completion both during class (about 44 min per day) and through the 4-month follow-up. Thus, there is compelling evidence that both classes were credible and engaging.

In short, our results suggest we were successful in demonstrating that HEP is an active control for MBSR that is inert with respect to mindfulness. These results are likely generalizable across different populations given the recruitment of a heterogeneous community sample and the similarity of the study's results and results from historical, non-study, MBSR classes.

There are several potential limitations of the study, including: (1) a possible weak MBSR program, (2) possible demand characteristics; (3) insufficient power to detect group differences on our PROs; (4) dosage effects; (5) intervention differences in the explicitness of pain regulation instruction; and (6) expectancy differences between HEP and MBSR.

First, it is possible that the lack of group differences in traditional PROs was due to an ineffective MBSR program, especially since the study lacked a wait-list control. However, our results indicate that the MBSR program was as effective as meta-analytic results for MBSR (e.g., Grossman, Niemann, Schmidt, & Walach, 2004) and results of a Monte Carlo study indicate that the current study's MBSR intervention is no less effective than historical MBSR interventions ($N = 534$) from the same UW-Health Program using the GSI or MSC. Furthermore, drop-out rates were similar to historical data from the same program.

Second, it is possible that demand characteristics account for the thermal pain rating results. To address this concern, we made the demand characteristic similar for both groups by introducing a within-subject manipulation with conditions relevant for each of the interventions. Nevertheless, demand characteristics may not have been comparable since the theme of pain regulation is more explicitly addressed in the MBSR than in the HEP intervention. Despite this, the differential intervention results may represent a promising improvement over commonly used PRO measures of mindfulness. However, further research is needed to support this possibility.

Third, our sample size was comparable to other MBSR studies and based on effect sizes reported in the literature. However, those effect sizes are based on studies that do not use an active control condition. Comparing MBSR to a well-designed control resulted in smaller effect sizes and therefore requires larger sample sizes to identify intervention differences. For example, the intervention effect for pre-post GSI change is $\eta^2 = .007$, corresponding to a intervention difference in change scores of .034 units on the GSI and a Cohen's $d = .17$ (small effect). Approximately 1400 participants per intervention are needed to achieve a power of .80 at $\alpha = .05$ for an effect this size and it is appropriate to ask whether such a small effect is worth pursuing. There were trend-level practice effects in our data that may have been significant with greater power. For example, one intervention \times time interaction for the GSI had an R^2 change = .05 corresponding to a Cohen's $d = .46$ (medium effect). Even an effect this size would require approximately 211 participants per intervention to achieve a power of .80 at an $\alpha = .05$. Effects such as the latter compare favorably to other treatment effects in psychotherapy research suggesting that they may be worth pursuing. We did have sufficient power to detect effects of time on various indices of mental and physical distress indicating that the interventions were effective at producing change. In sum, our primary null results are not likely due to power considerations.

Fourth, research indicates that the development of expertise in many endeavors requires intense practice of 1000 h or more (e.g., Ericsson, Krampe, & Tesch-Römer, 1993; Brefczynski-Lewis, Lutz, Schaefer, Levinson, & Davidson, 2007; Lutz, Greischar, Rawlings, Ricard, & Davidson, 2004; Lutz et al., 2009; Slagter et al., 2007 for expert meditators). It is thus possible that many benefits of mindfulness will not be evident at the dose delivered by an eight-week MBSR course (25 h in class + 31 h of practice outside of class between T1 and T2 = 56 h) or, indeed, the HEP class.

¹⁴ Due to the issues discussed in footnote #8 above, we again here note the limitations of our analyses in addressing growth across all three time points simultaneously, an issue that could be better addressed by including additional assessment points due to the nonlinear trajectories of change.

Fifth, although there is evidence that fitness and music training can moderate pain experiences (e.g., Tajet-Foxell & Rose, 1995; Siedliecki & Good, 2006; Zhao & Chen, 2009), HEP and MBSR differed in the explicitness with which pain regulation was addressed. MBSR focuses explicitly on pain regulation whereas HEP focused on reducing pain through the modification of class activities (e.g., if an activity was painful a modification was introduced) and through the benefits of the practices being taught (e.g., fitness).

Sixth, an ideal active control condition controls for expectancy effects. The fact that participants had a “somewhat strong” preference to be randomized to MBSR (64% preferred) over HEP (15% preferred; 21% had no preference) and also rated the intervention value as higher for MBSR than HEP provide evidence that HEP does not match MBSR in terms of expectancy or performance, potentially biasing the study against the HEP condition from the outset. Given the amount of marketing and cultural prominence of MBSR and mindfulness-related interventions, it is not surprising that participants had a preference to be randomized to MBSR. However, in contrast to this, there were no differences in class attendance between interventions, no differences in drop-out rates between groups, and actually more practice for HEP participants than MBSR participants. Furthermore, there was symptom improvement in both interventions over time but no group differences.

In conclusion, the lack of intervention differences on PROs often used to measure benefit in MBSR, combined with thermal pain evidence that mindfulness was present as an active ingredient in MBSR but not HEP, suggest that the HEP is a useful control condition for rigorous investigations of MBSR's relative efficacy when mindfulness is considered the active ingredient. Furthermore, although our results do not undermine the substantial evidence supporting the effectiveness of MBSR, they do suggest that the active ingredient of mindfulness in MBSR is no more effective than alternative active ingredients present in HEP for the PRO measures we employed. Future research that includes a wait-list or similar control would allow us to make more definitive comments about the efficacy of HEP for improving well-being. For now, we conclude that MBSR is as efficacious – but not more efficacious – than another active intervention (HEP) when applied to a typical MBSR population when our PROs are used. This conclusion represents an important shift in how we interpret the vast majority of MBSR outcomes in the extant literature. Furthermore, the fact that MBSR reduced ratings of thermal pain relative to the control condition, suggests that future research investigate whether the pain task may represent a promising measure of mindfulness. This suggests that future studies investigating mindfulness as a specific ingredient in MBSR (1) include control groups designed to address the questions being addressed, and (2) use behavioral or other more objective measures of intervention-specific skill acquisition in addition to PROs.

Acknowledgments

Donal MacCoon was responsible for developing the HEP intervention and various aspects of study design and analysis. Dr. MacCoon had full access to all the data in the study and takes responsibility for the integrity of the data and the accuracy of the data analysis. Zac Imel was a close collaborator on ensuring the HEP intervention design met the highest standards of scientific rigor from a common factors perspective and also helped with aspects of study design and analysis. Melissa Rosenkranz helped develop the HEP intervention and helped design and analyze different aspects of the study. Jenna Sheftel coordinated the study, playing a major role in maximizing participant retention and also inspired changes to the study design. Helen Weng designed various aspects of tasks used in the study. Jude Sullivan played a large role in designing and implementing HEP. Jude Sullivan reports that his salary and career

are based on work consistent with the HEP intervention and, thus, may represent a conflict of interest. Katherine Bonus played a large role in implementing MBSR and ensuring that study design did not compromise the delivery of MBSR. Katherine Bonus reports that her salary and career are based on work consistent with the MBSR intervention and, thus, may represent a conflict of interest. Catherine Stoney collaborated in study design and the design of HEP. Tim Salomons helped design the pain rating task and consulted on many aspects of its implementation. Richard Davidson was involved in supervising all aspects of the study. Antoine Lutz was the study's Primary Investigator, analyzed the pain rating data, and was integrally involved in all phases of study design and implementation. No author had any conflict of interest unless specified above.

This work was supported primarily by a grant from the National Center for Complementary and Alternative Medicine (U01AT002114-01A1 to Antoine Lutz), in addition to grants from the National Institute of Mental Health (P50-MH069315-03 to Richard Davidson), the Fetzer Institute, and gifts from Bryant Wangard, Ralph Robinson, Keith and Arlene Bronstein, John W. Kluge Foundation and the Impact Foundation (to Richard Davidson). HEP was developed in collaboration with Pam D. Young, M.A. (music therapist) and Julie P. Thurlow, DrPH, RD, CD (nutrition instructor) who are co-authors on the HEP Guidelines. Laura Lee Johnson, Ph.D. at the National Institutes of Health: National Center for Complementary & Alternative Medicine helped with HEP and study design. We are grateful for the participation of Cindy McCallum in co-teaching the MBSR interventions and for Laura Pinger's help with descriptions of MBSR. We also gratefully acknowledge facilitation and input from Helen Weng and Jackie Kuta-Bangsberg on several aspects of HEP. Saki Santorelli and Jon Kabat-Zinn from the Center for Mindfulness (CFM) at the University of Massachusetts Medical School graciously provided electronic copies of their MBSR guidelines and Jon Kabat-Zinn provided consultation. We also wish to thank Dan Bolt, John Curtin, Kristin Javares, Dana Tudorascu, and Alex Shackman for statistical consultations..

Appendix. Supplementary material

Supplementary data associated with this article can be found, in the online version, at [doi:10.1016/j.brat.2011.10.011](https://doi.org/10.1016/j.brat.2011.10.011).

References

- Barnhofer, T., Duggan, D., Crane, C., Hepburn, S., Fennell, M., & Williams, J. M. (2007). Effects of meditation on frontal alpha-asymmetry in previously suicidal individuals. *Neuroreport*, 7(18), 709–712.
- Baskin, T. W., Tierney, S. C., Minami, T., & Wampold, B. E. (2003). Establishing specificity in psychotherapy: a meta-analysis of structural equivalence of placebo controls. *Journal of Consulting and Clinical Psychology*, 71(6), 973–979.
- Bittman, B. B., Berk, L. S., Felten, D. L., Westengard, J., Simonton, O. C., Pappas, J., et al. (2001). Composite effects of group drumming music therapy on modulation of neuroendocrine-immune parameters in normal subjects. *Alternative Therapies*, 7(1), 38–47.
- Bittman, B., Berk, L., Shannon, M., Sharaf, M., Westengard, J., Guegler, K. J., et al. (2005). Recreational music-making modulates the human stress response: a preliminary individualized gene expression strategy. *Medical Science Monitor*, 11(2), 31–40.
- Bittman, B. B., Snyder, C., Bruhn, K. T., Liebfried, F., Stevens, C. K., Westengard, J., et al. (2004). Recreational music-making: an integrative group intervention for reducing burnout and improving mood states in first year associate degree nursing students: insights and economic impact. *International Journal of Nursing Education Scholarship*, 1(1), 1–26.
- Brefczynski-Lewis, J. A., Lutz, A., Schaefer, H. S., Levinson, D. B., & Davidson, R. J. (2007). Neural correlates of attentional expertise in long-term meditation practitioners. *PNAS Proceedings of the National Academy of Sciences of the United States of America*, 104(27), 11483–11488.
- Brown, C. A., & Jones, A. K. P. (2010). Meditation experience predicts less negative appraisal of pain: electrophysiological evidence for the involvement of anticipatory neural responses. *Pain*, 150(3), 428–438.

- Cohen-Katz, J., Wiley, S. D., Capuano, T., Baker, D. M., Kimmel, S., & Shapiro, S. (2005). The effects of mindfulness-based stress reduction on nurse stress and burnout, part II: a quantitative and qualitative study. *Holistic Nursing Practice*, 19(1), 26–35.
- Cook, R. D., & Weisberg, S. (1982). *Residuals and influence in regression*. New York: Chapman & Hall.
- Davidson, R. J., Kabat-Zinn, J., Schumacher, J., Rosenkranz, M., Muller, D., Santorelli, S. F., et al. (2003). Alterations in brain and immune function produced by mindfulness meditation. *Psychosomatic Medicine*, 65(4), 564–570.
- Derogatis, L. R. (1983). *SCL-90-R administration, scoring and procedures manual II*. Towson, MD: Clinical Psychometric Research.
- Ericsson, K. A., Krampe, R. T., & Tesch-Römer, C. (1993). The role of deliberate practice in the acquisition of expert performance. *Psychological Review*, 100(3), 363–406.
- Frank, J. D., & Frank, J. B. (1993). *Persuasion and healing: A comparative study of psychotherapy (paperback)*. Baltimore, MD US: Johns Hopkins University Press.
- Gaffan, E. A., Tsoussis, J., & Kemp-Wheeler, S. M. (1995). Researcher allegiance and meta-analysis: the case of cognitive therapy for depression. *Journal of Consulting and Clinical Psychology*, 63(6), 966–980.
- Gregg, J. A., Callaghan, G. M., Hayes, S. C., & Glenn-Lawson, J. L. (2007). Improving diabetes self-management through acceptance, mindfulness, and values: a randomized controlled trial. *Journal of Consulting and Clinical Psychology*, 75(2), 336–343.
- Grossman, P., Niemann, L., Schmidt, S., & Walach, H. (2004). Mindfulness-based stress reduction and health benefits: a meta-analysis. *Journal of Psychosomatic Research*, 57(1), 35–43.
- Grossman, P., Tiefenthaler-Gilmer, U., Raysz, A., & Kesper, U. (2007). Mindfulness training as an intervention for fibromyalgia: evidence of postintervention and 3-year follow-up benefits in well-being. *Psychotherapy and Psychosomatics*, 76(4), 226–233.
- Grunbaum, A. (1986). *Psychological Medicine*, 16(1), 19–38.
- Haefel, G. J., & Howard, G. S. (2010). Self-report: psychology's four-letter word. *The American Journal of Psychology*, 123(2), 181–188.
- Haskell, W. L., Lee, I., Pate, R. R., Powell, K. E., Blair, S. N., Franklin, B. A., et al. (2007). Physical activity and public health: updated recommendation for adults from the American college of sports medicine and the American heart association. *Medicine & Science in Sports & Exercise*, 1423–1434.
- Hollon, S. (1999). Allegiance effects in treatment research: a commentary. *Clinical Psychology: Science and Practice*, 6, 107–112.
- Hu, M. H., & Woollacott, M. H. (1994). Multisensory training of standing balance in older adults: II. kinematic and electromyographic postural responses. *Journal of Gerontology*, 49(2), M62–M71.
- Imel, Z., Baldwin, S., Bonus, K., & MacCoon, D. (2008). Beyond the individual: group effects in mindfulness-based stress reduction. *Psychotherapy Research*, 18(6), 735–742.
- Imel, Z. E., Wampold, B. E., Miller, S. D., & Fleming, R. R. (2008). Distinctions without a difference: direct comparisons of psychotherapies for alcohol use disorders. *Psychology of Addictive Behaviors*, 22(4), 533–543.
- Kabat-Zinn, J. (1982). An outpatient program in behavioral medicine for chronic pain patients based on the practice of mindfulness meditation: theoretical considerations and preliminary results. *General Hospital Psychiatry*, 4, 33–47.
- Kabat-Zinn, J. (1990). *Full catastrophe living: Using the wisdom of your body and mind to face stress, pain and illness*. New York: Delacorte.
- Kabat-Zinn, J., Wheeler, E., Light, T., Skillings, A., Scharf, M. J., Cropley, T. G., et al. (1998). Influence of a mindfulness meditation-based stress reduction intervention on rates of skin clearing in patients with moderate to severe psoriasis undergoing phototherapy (UVB) and photochemotherapy (PUVA). *Psychosomatic Medicine*, 60(5), 625–632.
- Kirsch, I. (2005). Placebo psychotherapy: synonym or oxymoron? *Journal of Clinical Psychology*, 61(7), 791–803.
- Kumar, A. M., Tims, F., Cruess, D. G., Mintzer, M. J., Ironson, G., Loewenstein, D., et al. (1999). Music therapy increases serum melatonin levels in patients with Alzheimer's disease. *Alternative Therapies*, 5(6), 49–57.
- Luborsky, L., Diguier, L., Luborsky, E., & Schmidt, K. A. (1999). The efficacy of dynamic versus other psychotherapies: is it true that 'everyone has won and all must have prizes?'—an update. In D. S. Janowsky (Ed.), *Psychotherapy indications and outcomes* (pp. 3–22). American Psychiatric Association.
- Lutz, A., Greischar, L. L., Rawlings, N. B., Ricard, M., & Davidson, R. J. (2004). Long-term meditators self-induce high-amplitude gamma synchrony during mental practice. *PNAS*, 101(46), 16369–16373.
- Lutz, A., Slagter, H. A., Rawlings, N. B., Francis, A. D., Greischar, L. L., & Davidson, R. J. (2009). Mental training enhances attentional stability: neural and behavioral evidence. *Journal of Neuroscience*, 29(42), 13418–13427.
- Ma, S. H., & Teasdale, J. D. (2004). Mindfulness-based cognitive therapy for depression: replication and exploration of differential relapse prevention effects. *Journal of Consulting and Clinical Psychology*, 72(1), 31–40.
- MacCoon, D. G., Sullivan, J. C., Davidson, R. J., Stoney, C. M., Christmas, P. D., Thurlow, J. P., et al. (2011). *Health-enhancement program (HEP) guidelines*. Retrieved from Permanent URL: <http://digital.library.wisc.edu/1793/28198>.
- McGuire, T. A., & Keene, J. S. (2006). The effect of a balance training program on the risk of ankle sprains in high school athletes. *The American Journal of Sports Medicine*, 34, 1103–1111.
- McMillan, T. M., Robertson, I. H., Brock, D., & Chorlton, L. (2002). Brief mindfulness training for attentional problems after traumatic brain injury: a randomised control treatment trial. *Neuropsychological Rehabilitation*, 12(2), 117–125.
- Mohr, D. C., Spring, B., Freedland, K. E., Beckner, V., Areal, P., Hollon, S. D., et al. (2009). The selection and design of control conditions for randomized controlled trials of psychological interventions. *Psychotherapy and Psychosomatics*, 78, 275–284.
- Nisbett, R. E., & Wilson, T. D. (1977). Telling more than we can know: verbal reports on mental processes. *Psychological Review*, 84(3), 231–259.
- Pace, T. W. W., Negi, L. T., Adame, D. D., Cole, S. P., Sivilli, T. I., Brown, T. D., et al. (2008). Effect of compassion meditation on neuroendocrine, innate immune and behavioral responses to psychosocial stress. *Psychoneuroendocrinology*, 34(1), 87–98.
- Perlman, D. M., Salomons, T. V., Davidson, R. J., & Lutz, A. (2010). Differential effects on pain intensity and unpleasantness of two meditation practices. *Emotion*, 10(1), 65–71.
- Pradhan, E. K., Baumgarten, M., Langenberg, P., Handwerker, B., Gilpin, A. K., Magyari, T., et al. (2007). Effect of mindfulness-based stress reduction in rheumatoid arthritis patients. *Arthritis and Rheumatism*, 57(7), 1134–1142.
- Salomons, T. V., Johnstone, T., Backonja, M. M., & Davidson, R. J. (2004). Perceived controllability modulates the neural response to pain. *Journal of Neuroscience*, 24(32), 7199–7203.
- Siedliecki, S. L., & Good, M. (2006). Effect of music on power, pain, depression and disability. *Journal of Advanced Nursing*, 54(5), 553–562.
- Slagter, H. A., Lutz, A., Greischar, L. L., Francis, A. D., Nieuwenhuis, S., Davis, J. M., et al. (2007). Mental training affects distribution of limited brain resources. *PLOS Biology*, 5(6), 1.
- Specia, M., Carlson, L. E., Goodey, E., & Angen, M. (2000). A randomized, wait-list controlled clinical trial: the effect of a mindfulness meditation-based stress reduction program on mood and symptoms of stress in cancer outpatients. *Psychosomatic Medicine*, 62(5), 613–622.
- Stevens, J. P. (1984). Outliers and influential data points in regression analysis. *Psychological Bulletin*, 95(2), 334–344.
- Tajet-Foxell, B., & Rose, F. D. (1995). Pain and pain tolerance in professional ballet dancers. *British Journal of Sports Medicine*, 29(1), 31–34.
- Thompson, C. (1989). Anxiety. In C. Thompson (Ed.), *The instruments of psychiatric research* (pp. 127–155). New York: Wiley.
- Travis, J. W. (1977). *Wellness workbook for health professionals*. Mill Valley, CA: Wellness Resource Center.
- U.S. Department of Health and Human Services and U.S. Department of Agriculture. (2005). *Dietary guidelines for Americans* (6th ed.). Washington, DC: U.S. Government Printing Office.
- Wachi, M., Koyama, M., Utsuyama, M., Bittman, B. B., Kitagawa, M., & Hirokawa, K. (2007). Recreational music-making modulates natural killer cell activity, cytokines, and mood states in corporate employees. *Medical Science Monitor*, 13(2), 57–70.
- Wampold, B. E. (2001). *The great psychotherapy debate: Models, methods, and findings*. Lawrence Erlbaum Associates Publishers.
- Wampold, B. E., Imel, Z. E., Laska, K. M., Benish, S., Miller, S. D., Flückiger, C., et al. (2010). Determining what works in the treatment of PTSD. *Clinical Psychology Review*, 30(8), 923–933.
- Wampold, B. E., Mondin, G. W., Moody, M., Stich, F., Benson, K., & Ahn, H. (1997). A meta-analysis of outcome studies comparing bona fide psychotherapies: empirically, 'all must have prizes'. *Psychological Bulletin*, 122(3), 203–215.
- Zhao, H., & Chen, A. C. N. (2009). Both happy and sad melodies modulate tonic human heat pain. *The Journal of Pain*, 10(9), 953–960.