

## SELF-REPORT MINDFULNESS VALIDITY

What can we learn from randomized clinical trials about the construct validity of self-report measures of mindfulness? A meta-analysis

Simon B. Goldberg<sup>\*1,2,3</sup>, Raymond P. Tucker<sup>4</sup>, Preston A. Greene<sup>3</sup>, Tracy L. Simpson, PhD<sup>3</sup>, William T. Hoyt<sup>1</sup>, David J. Kearney<sup>3</sup>, & Richard J. Davidson<sup>2,5</sup>

<sup>1</sup>Department of Counseling Psychology, University of Wisconsin - Madison, Madison, WI, USA

<sup>2</sup>Center for Healthy Minds, University of Wisconsin - Madison, Madison, WI, USA

<sup>3</sup>VA Puget Sound Health Care System, Seattle, Washington, USA

<sup>4</sup>Department of Psychology, Louisiana State University, Baton Rouge, LA, USA

<sup>5</sup>Department of Psychology, University of Wisconsin - Madison, Madison, WI, USA

\*Correspondence to: Simon B. Goldberg, Department of Counseling Psychology, University of Wisconsin – Madison, 335 Education Building, 1000 Bascom Mall, Madison, WI, 53706  
[sbgoldberg@wisc.edu](mailto:sbgoldberg@wisc.edu)

Citation: Goldberg, S. B., Tucker, R. P., Greene, P. A., Simpson, T.L., Hoyt, W. T., Kearney, D. J., & Davidson, R. J. (2019). What can we learn from randomized clinical trials about the construct validity of self-report measures of mindfulness? A meta-analysis. *Mindfulness*, 10(5), 775-785. doi: 10.1007/s12671-018-1032-y

Conflict of Interest: RJD is the founder, president, and serves on the board of directors for the non-profit organization, Healthy Minds Innovations, Inc. In addition, RJD serves on the board of directors for the Mind and Life Institute. The remaining authors (SG, RT, PG, TS, WH, DK) declare that they have no conflict of interest.

Compliance with Ethical Standards: Ethical approval: This article does not contain any studies with human participants or animals performed by any of the authors.

Funding: This work was funded by the National Center for Complementary and Alternative Medicine (Grant P01AT004952) and the Mind & Life Institute (Francisco J. Varela Award). NCCAM and Mind and Life Institute were not directly involved in study design; collection, analysis, or interpretation of data; in writing of the report; or in the decision to submit the article for publication. Any views, findings, conclusions, or recommendations expressed in this publication do not necessarily reflect those of the Mind & Life Institute.

### Abstract

Because they provide data on responsiveness to experimental manipulation, clinical trials involving mindfulness-based interventions are a source of evidence for the construct validity of self-report measures of mindfulness. Within-group and between-group changes in mindfulness were examined from randomized clinical trials comparing mindfulness interventions to other *bona fide* treatment comparison conditions or waitlist control conditions. We also examined changes in clinical outcomes and the magnitude of these changes relative to changes in mindfulness. We included 69 published studies representing 55 unique samples ( $n = 4,743$ ). Self-report mindfulness measures showed relatively larger gains in mindfulness intervention conditions vis-à-vis waitlist comparison conditions at both post-treatment (effect size [ES] = 0.52, 95% CI [0.40, 0.64]) and follow-up (ES = 0.52 [0.20, 0.84]), although the effect at follow-up diminished to non-significance in a trim-and-fill analysis intended to account for publication bias (ES = 0.35 [-0.03, 0.72]). Measures of mindfulness also showed relatively larger gains in mindfulness intervention conditions vis-à-vis *bona fide* comparison conditions, but only at post-treatment (ES = 0.25 [0.11, 0.38], 0.10 [-0.08, 0.28], at post-treatment and follow-up, respectively). All three conditions (mindfulness, *bona fide*, waitlist) showed relatively larger improvements on measures of clinical outcomes than measures of mindfulness, with the exception of waitlist conditions for which this effect was no longer significant at follow-up. Taken together, findings provide partial support for the unique responsiveness of mindfulness self-report measures to interventions that include promotion of mindfulness meditation practice.

*Keywords:* mindfulness; self-report measures; construct validity; meta-analysis

## SELF-REPORT MINDFULNESS VALIDITY

The construct of mindfulness is increasingly visible in psychology in recent decades. Mindfulness-based interventions, such as mindfulness-based stress reduction (MBSR; Kabat-Zinn, 1990) and mindfulness-based cognitive therapy (MBCT; Segal, Williams, & Teasdale, 2002) are being used to treat a wide variety of psychological and medical conditions (Goldberg et al., 2018; Goyal et al., 2014; Zoogman, Goldberg, Hoyt, & Miller, 2014). In addition, dispositional mindfulness has been associated with a host of psychological characteristics including psychiatric symptoms, wellbeing (Baer et al., 2008), and personality traits (Giluk, 2009), as well as with neurobiological and behavioral markers (Brown, Weinstein, & Creswell, 2013; Creswell, Way, Eisenberger, & Lieberman, 2007; Garland, Boettiger, Gaylord, Chanon, & Howard, 2011).

As mindfulness is incorporated into the psychological canon, it becomes vital that reliable and valid measures of this construct are available (Lutz, Jha, Dunne, & Saron, 2015). To date, several self-report measures of mindfulness have been developed. Two of the most popular measures of this kind are the Five Facet Mindfulness Questionnaire (FFMQ; Baer, Smith, Hopkins, Krietmeyer, & Toney, 2006) and the Mindful Attention Awareness Scale (MAAS; Brown & Ryan, 2003). Despite the widespread use of measures like the FFMQ and the MAAS, some have questioned the validity of self-report measures of mindfulness (Lutz et al., 2015). Among others, Grossman (2008) has raised several such concerns, calling for more rigorous assessment of these measures' psychometric properties. In particular, concerns have been raised regarding their construct validity (Goldberg et al., 2016; Van Dam et al., 2018), defined as the extent to which they measure what they are intended to measure (Crocker & Algina, 2008).

Construct validity inquiries seek to establish evidence that score variance reflects variance on the construct of interest, and to rule out that scores contain construct-irrelevant

## SELF-REPORT MINDFULNESS VALIDITY

variance (Crocker & Algina, 2008). Given most measures of mindfulness are self-report (although not all, e.g., Levinson, Stoll, Kindy, Merry, & Davidson, 2014), there are reasons to be skeptical about whether people accurately report their levels of mindfulness. If respondents are not generally aware or accurate in their self-perceptions (as is likely to be the case when an individual has a low level of mindfulness; Davidson & Kaszniak, 2015; Grossman, 2008), scores on the measures may instead reflect response biases such as social desirability (Tracey, 2016) or may reflect variance in conceptually distinct but psychologically related constructs (e.g., positive or negative mood).

One test of construct validity recommended by Cronbach and Meehl (1955) is to examine whether a measure behaves as predicted in response to experimental manipulation. Thus, a basic test of construct validity for mindfulness measures is responsiveness to experimental manipulations intended to enhance mindfulness. We define this tendency to change in response to experimental treatment as responsiveness. In a meta-analytic context, at a basic level we can ask whether the responsiveness for mindfulness-based interventions (comparing pre- and post-treatment means for participants in a mindfulness condition) differs significantly from zero.

Randomized clinical trials (RCTs) include other design features that invite more sophisticated tests of construct validity, especially RCTs testing mindfulness in clinical populations. Notably, RCTs involving mindfulness-based interventions conducted in clinical samples typically include both (a) comparison conditions and (b) measures of clinical outcomes. This suggests two additional critical tests of the validity of mindfulness self-reports in this experimental context: one that compares responsiveness within mindfulness measures and between conditions and one that derives an effect size reflecting comparative responsiveness between mindfulness measures and clinical outcome measures within conditions.

## SELF-REPORT MINDFULNESS VALIDITY

RCTs of mindfulness-based interventions include one or more comparison conditions, which allows assessment of relative responsiveness within mindfulness measures and between conditions. Broadly, comparison conditions can be classified as (a) specific active control conditions (i.e., *bona fide* treatments that are intended to be therapeutic; Wampold & Imel, 2015); (b) non-specific active controls (i.e., placebo treatments that are not intended to be therapeutic); or (c) waitlist controls. While *bona fide* comparison conditions can be defined by their inclusion of ingredients that are intended to be therapeutic, placebo control conditions can vary considerably from study to study (Baskin, Tierney, Minami, & Wampold, 2003), which makes comparisons with non-specific active controls difficult to interpret. In the current study, non-specific active controls ( $k = 4$ ) were excluded for this reason. *Bona fide* comparison conditions, however, provide an especially informative comparison, given they control for not only non-specific factors that contribute to efficacy of psychological treatments (Wampold & Imel, 2015) but they also include specific therapeutic techniques (such as challenging irrational beliefs, in the case of cognitive behavioral therapy). Waitlist control conditions provide no treatment (or in some cases treatment-as-usual), and are intended to control for history and maturation effects on the outcome variable (Shadish, Cook, & Campbell, 2001). By conducting meta-analyses using a mixture of non-mindfulness-based *bona fide* comparison conditions and waitlist control conditions, the effects of instruction in mindfulness can be experimentally isolated.

An initial test of the validity of mindfulness self-report measures examines whether responsiveness of mindfulness measures is significantly greater for participants exposed to a mindfulness-based intervention compared with those exposed to specific active controls or a waitlist control condition. Even though *bona fide* comparison conditions do not directly teach

## SELF-REPORT MINDFULNESS VALIDITY

mindfulness-enhancing techniques (e.g., mindfulness meditation), these treatments may target some features that could reasonably increase mindfulness (e.g., awareness of one's inner experience through cognitive behavioral therapy), thus we do not predict that changes in mindfulness will be absent in these conditions. Waitlist controls, in contrast, should not show increases in mindfulness over time.

RCTs conducted in clinical samples also typically include outcome measures targeted to the disorder under study. These measures can be used to assess the differential responsiveness between mindfulness measures and clinical outcomes, within conditions. For the RCTs considered here, all studies focused on some specific psychological problem (e.g., depression or anxiety) and included at least one measure of symptoms that characterize this problem (e.g., Beck Depression Inventory) along with a self-report measure of mindfulness. With multiple measures in each arm of the study, there is the possibility of examining the degree to which each type of outcome – measures of mindfulness and measures of clinical outcomes – is responsive to each of the two intervention types (i.e., to mindfulness or *bona fide* treatment interventions). We quantify the differential responsiveness of mindfulness measures and clinical outcome measures (in response to a particular experimental condition) as the difference between the effect size reflecting responsiveness of the mindfulness measure (expressed as a within-groups  $d$ , Becker, 1988) and that for the clinical outcome measure. Thus, differential responsiveness is conceptualized as a comparison between mindfulness measures and clinical outcomes within conditions.

In the *bona fide* comparison conditions the treatment targets psychological symptoms and any mindfulness effects are incidental; thus, for these treatments we expected the change on targeted symptom measures to exceed changes on measures of mindfulness. For the waitlist

## SELF-REPORT MINDFULNESS VALIDITY

control conditions, significant change was expected on neither the measures of mindfulness or clinical outcomes (although some regression to the mean can be expected on measures of clinical symptoms in clinical samples; Barnett, Van Der Pols, & Dobson, 2004). In the mindfulness-based treatment condition, we expected to see improvement on both the mindfulness measure and the targeted symptom measure, with no *a priori* expectations regarding which would increase more.

While concerns regarding the construct validity of mindfulness measures have been raised previously (Grossman, 2008; Goldberg et al., 2016; Van Dam et al., 2018), to our knowledge, no prior work has used meta-analytic methods to assess the discriminant validity of mindfulness measures using the differential responsiveness comparisons just described. However, prior RCTs and one meta-analysis assess between-group effects on measures of mindfulness (comparing relative responsiveness of mindfulness measures between conditions).

Using data from a RCT of MBSR, Goldberg et al. (2016) examined relative changes in FFMQ scores for participants assigned to MBSR, a *bona fide* comparison condition that was intended to be therapeutic (Health Enhancement Program [HEP]; MacCoon et al., 2012), or a waitlist condition. Goldberg et al. failed to find evidence for specific responsiveness to the mindfulness intervention: FFMQ scores demonstrated equivalent improvement over time for individuals receiving MBSR or HEP, with at least some of the FFMQ subscales showing larger gains in the MBSR and HEP conditions relative to the waitlist control.

A recent meta-analysis also examined the degree to which changes in measures of mindfulness (e.g., FFMQ, MAAS) were differentially influenced by experimental manipulation. Across 88 studies, Quaglia et al. (2016) found evidence suggesting mindfulness-based interventions produce larger changes in self-report measures of mindfulness relative to both

## SELF-REPORT MINDFULNESS VALIDITY

active and inactive (i.e., waitlist) control conditions across a range of mindfulness facets (i.e., attention, description, nonjudgment, nonreactivity, observation). In contrast to Goldberg et al. (2016), Quaglia et al.'s results support the notion that mindfulness measures show greater responsiveness to interventions involving mindfulness, compared with other active treatment conditions.

The aim of the present study is to establish whether self-report mindfulness measures are responsive to mindfulness interventions; whether they respond specifically to the mindfulness-enhancing techniques in these interventions (as opposed to factors common to other psychotherapeutic treatments); and whether they show discriminant validity from measures of psychological symptoms. Thus, we sought to extend Quaglia et al.'s (2016) findings by testing not only specificity of relative responsiveness to experimental manipulation (as examined by Quaglia et al.), but also differential responsiveness (i.e., discriminant validity) compared with measures of clinical outcomes. In order to evaluate differential responsiveness, we restricted our search to randomized trials of clinical interventions using clinical samples. In addition, we included as mindfulness treatments only interventions based on mindfulness meditation allowing assessment of a more homogeneous family of therapies (e.g., MBCT, MBSR), and excluded interventions (e.g., Acceptance and Commitment Therapy; Hayes, Strosahl, & Wilson, 1999) that are grounded in mindfulness theory but do not teach formal mindfulness meditation practices (i.e., sitting meditation). Finally, we examined changes in total scores rather than subscales of mindfulness measures, based on factor analytic evidence suggesting an overall mindfulness factor in commonly used measures of mindfulness (e.g., Baer et al., 2006; Brown & Ryan, 2003) and to reduce the number of analyses and increase the power of the statistical tests conducted.



## SELF-REPORT MINDFULNESS VALIDITY

Based on past findings, we made the following hypotheses. In regards to the relative responsiveness of mindfulness measures between conditions, we had three hypotheses. First ( $H_1$ ), given the focus of mindfulness-based interventions on the cultivation of mindfulness, we expected significant pre- to post-intervention and pre- to follow-up changes in mindfulness, for participants in the mindfulness conditions. Second ( $H_2$ ), we expected pre- to post-intervention and pre- to follow-up changes on mindfulness to be larger in the mindfulness condition, compared with alternative treatments and waitlist control conditions. However, many *bona fide* psychotherapeutic interventions emphasize mindfulness-relevant treatment elements such as introspection and self-awareness. Thus, for our third hypothesis ( $H_3$ ), we expected the mindfulness-to-waitlist comparison to be larger (reflecting greater changes in mindfulness scores) than the mindfulness-to-alternative-treatment comparison. In addition to assessing relative responsiveness, we derived differential responsiveness indices for each condition by subtracting the pre-post change effect size for the clinical outcome measure from that for the mindfulness measure. For our fourth hypothesis ( $H_4$ ), we expected differential responsiveness (within conditions) to be negative (greater change for the clinical outcome measure) in the alternative treatment condition and zero (no difference in responsiveness) in the waitlist control condition. We had no hypothesis regarding differential responsiveness in the mindfulness condition, as both mindfulness and clinical outcomes were expected to change in response to the treatment.

## Method

### Eligibility Criteria

We included RCTs of mindfulness-based interventions for adult patients with psychiatric and medical diagnoses that appear on the American Psychological Association's (APA) Division

## SELF-REPORT MINDFULNESS VALIDITY

12 (Society of Clinical Psychology; see Supplemental Materials Table 1, APA, 2017) list of disorders with known evidence-based treatments. To be eligible, samples had to have either a formal diagnosis or elevated symptoms of a given disorder. Samples receiving treatment within a facility focused on a specific disorder (e.g., substance abuse treatment) were included.

Elevated stress levels alone were not considered to reflect a clinical condition.

To qualify, the mindfulness interventions had to have mindfulness meditation as a core component with home meditation practice as a treatment ingredient. While interventions combining mindfulness with other modalities (e.g., mindfulness and cognitive techniques as in MBCT; Segal et al., 2002) were included, therapies emphasizing the attitudinal stance of mindfulness (rather than the formal practice of mindfulness meditation) were excluded (e.g., Acceptance and Commitment Therapy, Dialectical Behavior Therapy [DBT]; Hayes et al., 1999; Linehan, 1993). Other non-mindfulness forms of meditation (e.g., mantram repetition) were excluded. Interventions had to be delivered in real time (i.e., not provided through pre-recorded video instruction) and had to include more than one session (to allow for home meditation practice). Studies were also excluded for the following reasons: (1) not published in English; (2) not a peer-reviewed article; (3) data unavailable to compute standardized effect sizes (even after contacting study authors); (4) no disorder-specific (i.e., targeted) outcomes reported; (5) no measure of mindfulness included; (6) data redundant with other included studies; (7) no non-mindfulness-based intervention or condition included (i.e., the trial compared only two or more mindfulness-based interventions); (8) no waitlist (or TAU that was provided to both the mindfulness and control condition) or *bona fide* comparison condition included.

### Information sources

## SELF-REPORT MINDFULNESS VALIDITY

We searched the following databases: PubMed, PsycInfo, Scopus, Web of Science. In addition, a publically available comprehensive repository of mindfulness studies that is updated monthly was also searched (Black, 2012). Citations from recent meta-analyses and systematic reviews were also included. Citations were included from the first available date (i.e., 1966) until January 2<sup>nd</sup>, 2017.

### **Search**

We used the search terms “mindfulness” and “random\*”. When a database allowed (e.g., PsycInfo), we restricted our search to clinical trials.

### **Study Selection**

Titles and/or abstracts of potential studies were independently coded by the first author and a second co-author. Disagreements were discussed with a senior author until a consensus was reached.

### **Data Collection Process**

Standardized spreadsheets were developed for coding both study-level and effect size-level data. Coders were trained by the first author through coding an initial sample of studies ( $k = 10$ ) in order to achieve reliability. Data were extracted independently by the first author and a second co-author. Disagreements were discussed with a senior author. Inter-rater reliabilities were in the good to excellent range (Cicchetti, 1994):  $Ks > .60$  and  $ICCs > .80$  in the current study. When sufficient data for computing standardized effect sizes were unavailable, study authors were contacted.

### **Data Items**

Along with data necessary for computing standardized effect sizes, the following data were extracted: (1) disorder; (2) intent-to-treat (ITT) sample size; (3) whether an ITT analysis

## SELF-REPORT MINDFULNESS VALIDITY

was reported; (4) sample demographics (mean age, percentage female, percentage with some college education); (5) country of origin; (6) type of comparison condition.

Type of comparison condition was coded based on a two-tier system: waitlist conditions and *bona fide* comparison conditions. Waitlist conditions included waitlist controls as well as treatment-as-usual (TAU) conditions in which both the mindfulness and non-mindfulness arms received this treatment (i.e., there was no additional treatment provided to the TAU group). The *bona fide* treatment conditions included comparisons that were based on actual therapies and included specific treatment ingredients and mechanisms of change (Wampold & Imel, 2015). The decision to code using this scheme was based on evidence that whether a comparison group represents a *bona fide* comparison condition significantly influences the relative efficacy of mindfulness-based interventions (Goldberg et al., 2018). Some studies included both *bona fide* and waitlist comparison condition ( $k = 8$ ). In order to avoid duplicated data (i.e., comparing the mindfulness condition to both controls), we included only the *bona fide* comparison condition in between-group analyses.

### **Risk of Bias in Individual Studies**

Considerations for minimizing bias in individual studies were drawn from both Jadad's criteria (Jadad et al., 1996) as well as the GRADE system (Atkins et al., 2004). Based on the GRADE recommendation to select relevant study characteristics to quantify (Agency for Healthcare Research and Quality, 2014) and based on the large number of potential study characteristics for assessing quality in psychotherapy trials (e.g.,  $n = 185$  quality criteria; Liebherz, Schmidt, & Rabung, 2016), we restricted our analysis to randomized trials, employed intent-to-treat samples (when available), and coded the strength of the comparison condition.

### **Effect Size Computation**

## SELF-REPORT MINDFULNESS VALIDITY

For each research hypothesis, we developed an effect size for the comparison of interest as described below. When multiple outcomes of the same type (mindfulness or clinical symptoms) were included in the same study, data were aggregated within-studies using the ‘MAJ’ package (Del Re & Hoyt, 2010), following procedures described in Borenstein, Hedges, Higgins, & Rothstein (2009).

**Effect size calculation for relative responsiveness hypotheses.** We quantified responsiveness of mindfulness scores (within conditions) by computing  $d_{within}$  for each experimental condition.

$$d_{within} = \frac{M_{post} - M_{pre}}{SD_{pooled}} \quad (1a)$$

$$var(d_{within}) = \left( \frac{1}{n} + \frac{d^2}{2n} \right) \cdot 2(1 - r), \quad (1b)$$

Where  $r$  is the correlation between pre- and post-scores on mindfulness. As is typically the case in meta-analyses of clinical trials, the primary studies did not report  $r$ , so we imputed a correlation of  $r_{XX} = .50$  between time points (somewhat lower than a typical test-retest correlation, to account for intervention effects; see Hoyt & Del Re, in press). These effect sizes were corrected for bias, converting to Hedges’  $g_{within}$  as recommended by Borenstein, Hedges, Higgins, & Rothstein (2009). Within-condition effect sizes were computed from pre- to post-treatment (or time point closest to post-treatment) as well as from pre- to last available follow-up time point.

We then quantified relative responsiveness (to the mindfulness intervention compared with the two comparison conditions) as the difference in the pre-post effects (i.e., change scores). The resulting effect size (called  $\Delta$ , following Becker, 1988) represents the amount by which change in mindfulness in the mindfulness condition exceeds change in mindfulness in the comparison condition, in standard deviation units.

## SELF-REPORT MINDFULNESS VALIDITY

$$\Delta = g_{within}^M - g_{within}^C \quad (2a)$$

$$var(\Delta) = var(g_{within}^M) + var(g_{within}^C), \quad (2b)$$

where the  $M$  and  $C$  superscripts refer to the mindfulness and comparison conditions, respectively.

**Effect size calculation for differential responsiveness hypotheses.** In the second set of hypotheses, we quantified differential responsiveness (i.e., for the mindfulness measure compared with the clinical outcome measure) by computing a dependent samples  $\Delta_{dep}$  for each condition. Because this effect size is a difference between dependent estimates (i.e., two estimates derived from the same sample), the variance formula needs to take into account the correlation between the mindfulness and the clinical symptom effect sizes.

$$\Delta_{dep} = g_{within}^{mindful} - g_{within}^{clinical} \quad (3a)$$

$$var(\Delta_{dep}) = var(g_{within}^{mindful}) + var(g_{within}^{clinical}) - 2 \cdot r \cdot \sqrt{g_{within}^{mindful} \cdot g_{within}^{clinical}} \quad (3b)$$

Correlations between mindfulness and clinical measures were often not reported in the primary studies. Consequently, we used an imputed value of  $r = .50$ , based on meta-analytic estimates of the association between dispositional mindfulness and neuroticism (Giluk, 2009). (The sign of the correlation coefficient is positive because we reversed-scored clinical outcomes, so that positive effect sizes indicate improvement over time for both outcome variables.) Positive values of  $\Delta_{dep}$  reflect greater responsiveness for the mindfulness measure (compared to the clinical measure) in the condition under study.

Analyses were conducted using the R statistical software and the ‘metafor’ and ‘MAd’ packages (Del Re & Hoyt, 2010; Viechtbauer, 2010). Random effects models were used with a

## SELF-REPORT MINDFULNESS VALIDITY

restricted maximum-likelihood estimator and were weighted based on the inverse of the variance. Heterogeneity was assessed using the  $Q$ -statistic and quantified using  $I^2$ .

### **Risk of Bias Across Studies**

We assessed publication bias by visually inspecting funnel plots for asymmetry within the comparison of interest. In addition, primary models were re-estimated using trim-and-fill methods that account for the asymmetric distribution of studies around an omnibus effect (Viechtbauer, 2010).

## **Results**

### **Study Selection**

A total of 9,067 citations were retrieved. After 3,485 duplicates were removed, 5,582 unique titles and/or abstracts were coded. Following the application of the exclusion criteria (see flow diagram in Supplemental Materials), 69 articles including 55 unique samples were retained for analysis representing 4,743 participants.

### **Study Characteristics**

Effect sizes in standardized units (i.e.,  $d$ ) reflecting within-group and between-group changes on mindfulness as well as the relative responsiveness of mindfulness and clinical outcomes are shown in the Supplemental Materials along with other study characteristics (Table 2). The sample was on average 44.20 years old, 61.48% female, with 63.67% having some post-secondary education. The largest percentage of trials was conducted in the United States (52.73%). Approximately half of studies included waitlist control conditions (45.45%) and half included *bona fide* comparison conditions (54.55%). The most commonly studied disorder was depression (23.64%), followed by pain (21.82%), anxiety (16.36%), and addiction (9.09%). The majority of studies (58.18%) used either the FFMQ or the Kentucky Inventory of Mindfulness

## SELF-REPORT MINDFULNESS VALIDITY

Skills (KIMS; Baer, Smith, & Allen, 2004) to assess self-reported mindfulness; another 18.18% used the MAAS (with one study including both the FFMQ and the MAAS); the remaining studies ( $k = 12$ ) used other self-report mindfulness measures.

### **Risk of Bias Within Studies**

All included studies used randomized designs. More than half of the studies reported at least one ITT analysis (63.64%). When available, results from the ITT analysis were used.

### **Results of Individual Studies**

For each included study, treatment effects on self-report measures of mindfulness and clinical outcomes are reported in Supplemental Materials.

### **Mindfulness Measures: Responsiveness to Intervention**

The top half of Table 1 shows pre- to post-intervention and pre- to follow-up effect sizes by condition, for both mindfulness and clinical outcome measures. As expected ( $H_1$ ), there was evidence of significant changes in self-reported mindfulness in response to mindfulness interventions ( $g = 0.49$  [0.39, 0.58] from pre- to post-treatment;  $g = 0.31$  [0.17, 0.45] from pre- to follow-up). The parallel effect sizes for mindfulness responsiveness were close to zero (and not significantly different from zero) in the waitlist conditions, and were intermediate (and significantly different from zero) in the alternative treatment conditions.

### **Relative responsiveness of mindfulness measures across experimental conditions.**

The top half of Table 2 summarizes effect sizes (Becker's  $\Delta$ ) comparing responsiveness in mindfulness scores between conditions (see Supplemental Materials for forest plots). As expected ( $H_2$ ) mindfulness measures demonstrated enhanced responsiveness to mindfulness-based interventions relative to waitlist controls ( $\Delta = 0.52$ , [0.40, 0.64] pre- to post-treatment;  $\Delta = 0.52$ , [0.20, 0.84] from pre- to follow-up), and also relative to alternative, non-mindfulness-based



## SELF-REPORT MINDFULNESS VALIDITY

*bona fide* comparison conditions ( $\Delta = 0.25$ , [0.11, 0.38] pre- to post-treatment); however, the latter comparison was no longer significant at follow-up ( $\Delta = 0.10$ , [-0.08, 0.28]). Also in accordance with our predictions ( $H_3$ ), responsiveness effect sizes relative to waitlist conditions were larger than those relative to *bona fide* treatment comparisons at both time points ( $p < .05$ ) (although the robustness of the follow-up finding was called into question in the sensitivity analysis, as discussed in the later section on risk of bias).

### **Differential responsiveness between mindfulness measures and clinical outcomes.**

Our final set of hypotheses examined discriminant validity of mindfulness measures and clinical outcome measures in the context of experimental manipulation. Differential responsiveness effect sizes were computed within conditions as the difference between within-group  $d$ s for mindfulness and clinical outcome measures ( $\Delta_{dep}$ ), then meta-analyzed across studies, with the results summarized in Table 3. We predicted ( $H_4$ ) that differential responsiveness would be negative (reflecting greater responsiveness for the clinical outcome measure) in the alternative treatment condition and near zero for the waitlist condition. We made no prediction regarding whether clinical outcomes or measures of mindfulness would change more in the mindfulness conditions.

As shown in Table 3, we found negative differential responsiveness (i.e., change in mindfulness was smaller than change in clinical symptoms) in all three conditions. This difference in responsiveness was statistically significant (i.e., 95% CI excluded zero) for five of the six tests (three conditions; post-treatment and follow-up comparisons) except the test of the change to follow-up for the waitlist ( $\Delta_{dep} = -0.32$  [-0.65, 0.01]), which had the smallest amount of data available ( $k = 8$ ), and therefore the lowest statistical precision (and power). This result supported our prediction for *bona fide* comparison conditions, although the negative differential

## SELF-REPORT MINDFULNESS VALIDITY

responsiveness was not predicted in the waitlist condition. We consider possible explanations for this unexpected finding in the Discussion section.

### **Risk of Bias Across Studies**

Bias in the above analyses was assessed through funnel plots and trim-and-fill analyses. Asymmetric funnel plots suggested evidence for publication bias for several models (see Supplemental Materials for funnel plots). Trim-and-fill analyses yielded adjusted effect sizes, although the direction of adjustment varied (i.e., some effects became larger). The sensitivity analyses called into question one effect that appeared significant in the main analyses: pre- to follow-up between-group relative responsiveness on mindfulness measures in mindfulness versus waitlist control conditions (adjusted  $\Delta = 0.35$ ,  $[-0.03, 0.72]$ ; Table 2).

### **Discussion**

Our goal in this study was to examine evidence for construct validity of self-report measures of mindfulness derived from clinical trials that included a mindfulness intervention condition. These RCTs allow for robust examination of responsiveness to experimental manipulation, as described by Cronbach and Meehl (1955). Our meta-analytic findings provided support for the predictions ( $H_1$  to  $H_3$ ) that scores on mindfulness measures are responsive to experimental intervention: These measures registered moderate amounts of change in response to mindfulness interventions, little or no change in waitlist conditions, and intermediate levels of change in conditions implementing a non-mindfulness based alternative treatment.

While these results mirror those of previous reports (Quaglia et al., 2016), it is worth noting explicitly here that patients report changes in mindfulness in both mindfulness and non-mindfulness-based interventions (albeit to a smaller degree in non-mindfulness-based interventions). Changes in mindfulness induced by non-mindfulness-based interventions could

## SELF-REPORT MINDFULNESS VALIDITY

be due to a number of factors. This effect might indicate that the non-mindfulness-based interventions are implicitly or explicitly teaching mindfulness skills (e.g., meta-cognitive skills in the case of cognitive behavioral therapy). Alternatively, the responsiveness of mindfulness measures to non-mindfulness interventions may reflect construct-irrelevant variance (Hoyt, Warbasse, & Chu, 2006), such as general negative affect, that contributes to variance in mindfulness scores—a limitation in the construct validity of self-report measures of mindfulness (Goldberg et al., 2016; Grossman, 2008). Further research examining measures of mindfulness in the context of non-mindfulness-based interventions, as well as research employing multimethod assessment of mindfulness, can be helpful for clarifying what sources of variance contribute to scores on self-report measures of mindfulness (cf. Cronbach & Meehl, 1955).

A second set of hypotheses examined differential responsiveness of mindfulness and clinical outcome measures. These analyses used meta-analytic methods to examine a type of discriminant validity in the experimental context. We predicted ( $H_4$ ) that responsiveness (i.e., change) for mindfulness measures should be smaller than responsiveness of clinical outcome measures in the *bona fide* (non-mindfulness) intervention condition and should be similar (and near zero) in the waitlist control condition. Given that we expected change on both measures of mindfulness and measures of clinical outcomes in the mindfulness condition, no hypothesis was made about differential responsiveness in this group.

Of our two directional hypotheses, only the hypothesis relating to *bona fide* comparison conditions was supported. As predicted, changes in clinical outcomes exceeded those of changes in measures of mindfulness, supporting the prediction of discriminant responsiveness to *bona fide*, non-mindfulness based mental health interventions.

## SELF-REPORT MINDFULNESS VALIDITY

Interestingly, the same pattern was observed for the waitlist and mindfulness comparisons as well. The presence of relatively larger effects on clinical outcomes than measures of mindfulness in the waitlist condition underscores a challenge for differential responsiveness predictions based on clinical trials data: the possibility of differential improvement in the absence of treatment. Although we predicted equivalent (and near-zero) improvement for both sets of outcomes in the waitlist condition, there are at least three reasons that one might expect clinical symptoms to improve in the waitlist condition: regression to the mean; benefits of “treatment-as-usual” (given that it is generally not possible to prohibit control group participants from seeking assistance outside the study); and remoralization effects of the decision to seek treatment through participating in a research study (which may include seeking non-professional support and taking other actions outside the treatment context to ameliorate symptoms).

The presence of relatively larger effects on clinical outcomes than measures of mindfulness in the mindfulness condition is intriguing. While we did not have an *a priori* hypothesis related to this comparison, it is notable that the effect of mindfulness-based interventions on clinical outcomes is larger than that observed on measures of mindfulness, one of the key putative mediators of treatment effects in mindfulness interventions (Gu, Strauss, Bond, & Cavanagh, 2015). In theory, one might expect effects on mediators to be similar or larger than effects on clinical outcomes, because the intervention is the proximal cause of the mediator variable, and a distal cause (to the extent that the mediator explains the relation between intervention and outcome) of symptom reduction. Indeed, there is a strong consensus among mediation researchers that it is reasonable to search for mediated (indirect) effects even in the absence of a bivariate relation between the predictor variable and the outcome (Kenny, Kashy, & Bolger, 1998; MacKinnon, 2008; Shrout and Bolger, 2002), which reinforces the

## SELF-REPORT MINDFULNESS VALIDITY

notion that relations between the predictor and mediator may often be more robust than those between the predictor and outcome (the “total effect” in mediator models; Baron & Kenny, 1986; MacKinnon, 2008). In, their meta-analysis of mindfulness as a mediator in mindfulness-based interventions, Gu et al. (2015) reported that intervention effects on mindfulness were somewhat larger than those on clinical outcome ( $r_s = .34$  and  $.27$ , for effects on mindfulness and clinical outcomes, respectively). Our finding of a small but statistically significant difference in effect size favoring the clinical outcome measures may be attributable to the restriction of our review to clinical samples, and likely reflects additional pathways (i.e., beyond the mediated effect through changes in mindfulness) by which mindfulness-based interventions induce reductions in clinical symptoms (e.g., therapeutic alliance; Goldberg, Davis, & Hoyt, 2013).

### **Limitations**

Several limitations are worth acknowledging. The first is that our results were limited to published studies. Given the extensive nature of our literature search, we chose to exclude unpublished studies. However, publication bias is an increasing concern in psychology (DeCoster, Sparks, Sparks, Sparks, & Sparks, 2015), and our sensitivity analyses (trim-and-fill, funnel plots) suggest the presence of publication bias in our sample. As null results have historically been more difficult to publish (or have been intentionally omitted from published studies; DeCoster et al., 2015), it is likely that the treatment differences we observed on self-report measures of mindfulness overestimate the true differences. A second limitation was not disaggregating by mindfulness component (i.e., measure or subscale). This was done to limit the number of analyses and increase statistical power, but may have impacted of ability to detect differences in measure performance across specific aspects of mindfulness. A third limitation was not separating analyses by disorder. This would have allowed assessment of the extent to

## SELF-REPORT MINDFULNESS VALIDITY

which changes in mindfulness compared with changes in outcomes for different disorders. We chose not to explore this possibility due to the small number of certain disorder types (e.g., ADHD), particularly when crossed with comparison group type. Future studies, presumably using trials that are yet to be published, could explore some of these possibilities. A final limitation was the possibility of limited statistical power, particularly for certain analyses (e.g., those involving comparisons with waitlist conditions at follow-up). It is conceivable that certain effects were not detected due to Type II error.

Taken together, results from the current study provide partial support for the construct validity of self-report measures of mindfulness. Although responsive to mindfulness training, these measures appear to also change through other *bona fide* treatments, albeit to a lesser degree. Effects of mindfulness interventions on measures of mindfulness are also smaller than their effects on targeted outcomes, at least within the clinical samples included here.

As Cronbach and Meehl (1955) point out, instances of uncertain construct validity could implicate the measures used and/or the theory underlying the measures. This underscores the value in continued work on the measurement of mindfulness as well as efforts to untangle the mechanisms at play in mindfulness interventions. Future studies of mindfulness-based interventions will ideally include behavioral and neurobiological assessment of mindfulness and characteristics putatively related to mindfulness, along with self-report measures of mindfulness. Results from RCTs using these measures, particularly when also using comparison conditions that are intended to be therapeutic (Goldberg et al., 2017), can help assess the degree to which specific effects related to training in mindfulness are present. The development of novel assessment methods (e.g., significant-other ratings, observer ratings, mindfulness teacher ratings) may provide valuable alternatives to self-report measures of mindfulness in future studies.

## References

- Agency for Healthcare Research and Quality. (2014). *Methods guide for effectiveness and comparative effectiveness reviews*. Rockville, MD: Agency for Healthcare Research and Quality.
- American Psychological Association. (2017, January 2). *Research-supported psychological treatments*. Retrieved from: <https://www.div12.org/psychological-treatments/>
- Atkins, D., Eccles, M., Flottorp, S., Guyatt, G.H., Henry, D., Hill, S.,...& The GRADE Working Group. (2003). Systems for grading the quality of evidence and the strength of recommendations I: Critical appraisal of existing approaches The GRADE Working Group. *BMC Health Services Research*, 4(38). doi:10.1186/1472-6963-4-38
- Baer, R. A., Smith, G. T., & Allen, K. B. (2004). *Assessment of mindfulness by self-report: The Kentucky Inventory of Mindfulness Skills*. *Assessment*, 11, 191-206.
- Baer, R.A., Smith, G.T., Hopkins, J., Krietemeyer, J., & Toney, L. (2006). Using self-report assessment methods to explore facets of mindfulness. *Assessment*, 13, 27–45.
- Baer, R., Smith, G., Lykins, E., Button, D., Krietemeyer, J., Sauer, S....Williams, M.G. (2008). Construct validity of the Five Facet Mindfulness Questionnaire in meditating and nonmeditating samples. *Assessment*, 15, 329–342. doi: 10.1177/1073191107313003
- Barnett, A. G., Van Der Pols, J. C., & Dobson, A. J. (2004). Regression to the mean: what it is and how to deal with it. *International Journal of Epidemiology*, 34(1), 215-220.
- Baskin, T.W., Tierney, S.C., Minami, T., & Wampold, B.E. (2003). Establishing specific in psychotherapy: A meta-analysis of structural equivalence of placebo controls. *Journal of Consulting and Clinical Psychology*, 71(6), 973-979. doi: 10.1037/0022-006X.71.6.973
- Becker, B. (1988). Synthesizing standardized mean-change measures. *British Journal of Mathematical and Statistical Psychology*, 41, 257-278.
- Black, D.S. (2012). Mindfulness research guide: A new paradigm for managing empirical health information. *Mindfulness*, 1(3), 174-176. doi:10.1007/s12671-010-0019-0
- Borenstein, M., Hedges, L. V., Higgins, J. P. T., & Rothstein, H. R. (2009). *Introduction to meta-analysis*. New York: Wiley.
- 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.
- Brown, K.W., Weinstein, N., & Creswell, J.D. (2013). Trait mindfulness modulates neuroendocrine and affective responses to social evaluative threat. *Psychoneuroendocrinology*, 37, 2037-2041. doi: 10.1016/j.psyneuen.2012.04.003
- Cicchetti, D. (1994). Guidelines, criteria, and rules of thumb for evaluating normed and standardized assessment instruments in psychology. *Psychological Assessment*, 6(4), 284-290.
- Crocker, L., & Algina, J. (2008). *Introduction to classical and modern test theory*. Mason, Ohio: Cengage Learning.
- Creswell, J.D., Way, B.M., Eisenberger, N.I., & Lieberman, M.D. (2007). Neural correlates of dispositional mindfulness during affect labeling. *Psychosomatic Medicine*, 69, 560-565.
- Cronbach, L.J., & Meehl, P.E. (1955). Construct validity in psychological tests. *Psychological Bulletin*, 52(4), 281-302.
- Davidson, R.J., & Kaszniak, A.W. (2015). Conceptual and methodological issues in research on mindfulness and meditation. *American Psychologist*, 70(7), 581-592. doi: 10.1037/a0039512

## SELF-REPORT MINDFULNESS VALIDITY

- DeCoster, J., Sparks, E.A., Sparks, J.C., Sparks, G.G., & Sparks, C.W. (2015). Opportunistic biases: Their origins, effects, and an integrated solution. *American Psychologist*, 70(6), 499-514. doi: 10.1037/a0039191
- Del Re, A. C., Hoyt, W. T. (2010). *MAd: Meta-analysis with mean differences*. R package version 0.8, <http://CRAN.R-project.org/package=MAd>
- Garland, E.L., Boettiger, C.A., Gaylord, S., Chanon, V.W., & Howard, M.O. (2011). Mindfulness is inversely associated with alcohol attentional bias among recovering alcohol-dependent adults. *Cognitive Therapy and Research*, 36, 441-450.
- Giluk, T.L. (2009). Mindfulness, big five personality, and affect: A meta-analysis. *Personality and Individual Differences*, 47, 805-811.
- Goldberg, S.B., Davis, J.M., & Hoyt, W.T. (2013). The role of therapeutic alliance in mindfulness interventions: Therapeutic alliance in Mindfulness Training for Smokers. *Journal of Clinical Psychology*, 69(9), 936-950. doi: 10.1002/jclp.21973
- Goldberg, S.B., Tucker, R.P., Greene, P.A., Davidson, R.J., Wampold, B.E., Kearney, D.J., & Simpson, T.L. (2018). Mindfulness-based interventions for psychiatric disorders: A meta-analysis. *Clinical Psychology Review*, 59, 52-60. doi:10.1016/j.cpr.2017.10.011
- Goldberg, S. B., Tucker, R. P., Greene, P. A., Simpson, T.L., Kearney, D. J., & Davidson, R. J. (2017). Is mindfulness research methodology improving over time? A systematic review. *PLoS ONE*, 12(10), e0187298. doi: 10.1371/journal.pone.0187298
- Goldberg, S.B., Wielgosz, J., Dahl, C., Schuyler, B., MacCoon, D.S., Rosenkranz, M.,...& Davidson, R.J. (2016). Does the Five Facet Mindfulness Questionnaire measure what we think it does? Construct validity evidence from an active controlled randomized clinical trial. *Psychological Assessment*, 28(8), 1009-1014. doi: 10.1037/pas0000233
- Goyal, M., Singh, S., Sibinga, E.M., Gould, N.F., Rowland-Seymour, A., Sharma, R.,...& Haythornthwaite, J.A. (2014). Meditation programs for psychological stress and well-being: A systematic review and meta-analysis. *JAMA Internal Medicine*, 174(3), 357-368.  
doi:10.1001/jamainternmed.2013.13018
- Grossman, P. (2008). On measuring mindfulness in psychosomatic and psychological research. *Journal of Psychosomatic Research*, 64, 405-408.
- 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
- Hayes, S.C., Strosahl, K., & Wilson, K. (1999). *Acceptance and commitment therapy: An experiential approach to behavior change*. New York: Guilford Press.
- Hoyt, W. T., & Del Re, A. C. (2018). Effect size calculation in meta-analyses of psychotherapy outcome research. *Psychotherapy Research*, 28, 379-388.  
doi: 10.1080/10503307.2017.1405171
- Hoyt, W., Warbasse, R., & Chu, E. (2006). Construct validation in counseling psychology research. *The Counseling Psychologist*, 34(6), 769-805.
- Jadad, A.R., Moore, A., Carroll, D., Jenkinson, C., Reynolds, D.J.M., Gavaghan, D.J., et al. (1996). Assessing the quality of reports of randomized clinical trials: Is blinding necessary? *Controlled Clinical Trials*, 17, 1-12.
- Kabat-Zinn, J. (1990). *Full catastrophe living: Using the wisdom of your body and mind to face stress, pain, and illness*. New York: Delta.



## SELF-REPORT MINDFULNESS VALIDITY

- Levinson, D.B., Stoll, E.L., Kindy, S.D., Merry, H.L., & Davidson, R.J. (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
- Liebherz, S., Schmidt, N., & Rabung, S. (2016). How to assess the quality of psychotherapy outcome studies: A systematic review of quality assessment criteria. *Psychotherapy Research*, 26(5), 573-589. doi: 10.1080/10503307.2015.1044763
- Linehan, M.M. (1993). *Cognitive-behavioral treatment of borderline personality disorder*. New York: Guilford Press.
- Lutz, A., Jha, A. P., Dunne, J. D., & Saron, C. D. (2015). Investigating the phenomenological matrix of mindfulness-related practices from a neurocognitive perspective. *American Psychologist*, 70(7), 632-658. doi: 10.1037/a0039585
- MacCoon, D.G., Imel, Z.E., Rosenkranz, M.A., Sheftel, J.G., Weng, H.Y., Sullivan, J.C.,...Lutz, A. (2012). The validation of an active control intervention for Mindfulness Based Stress Reduction (MBSR). *Behaviour Research and Therapy*, 50, 3-12.
- Quaglia, J. T., Braun, S. E., Freeman, S. P., McDaniel, M. A., & Brown, K. W. (2016). Meta-analytic evidence for effects of mindfulness training on dimensions of self-reported dispositional mindfulness. *Psychological Assessment*, 28(7), 803-818. doi: 10.1037/pas0000268
- Segal, Z., Williams, J.W., & Teasdale, J. (2002). *Mindfulness-based cognitive therapy for depression: A new approach to preventing relapse*. New York: Guilford Press.
- Shadish, W.R., Cook, T.D., & Campbell, D.T. (2002). *Experimental and quasi-experimental designs for generalized causal inference*. Boston: Houghton Mifflin.
- Shrout, P.E., & Bolger, N. (2002). Mediation in experimental and nonexperimental studies: New procedures and recommendations. *Psychological Methods*, 7(4), 422-445.
- Tracey, T.J.G. (2016). A note on socially desirable responding. *Journal of Counseling Psychology*, 63(2), 224-232. doi: 10.1037/cou0000135
- Van Dam, N. T., van Vugt, M. K., Vago, D. R., Schmalzl, L., Saron, C. D., Olendzki, A., ... & Meyer, D. E. (2018). Mind the hype: A critical evaluation and prescriptive agenda for research on mindfulness and meditation. *Perspectives on Psychological Science*, 13(1), 36-61. doi: 10.1177/1745691617709589
- Viechtbauer, W. (2010). Conducting meta-analyses in R with the metafor package. *Journal of Statistical Software*, 36(3), 1-49.
- Wampold, B., & Imel, Z.E. (2015). *The great psychotherapy debate: The evidence for what makes psychotherapy work (2nd ed.)*. New York: Routledge.
- Zoogman, S., Goldberg, S.B., Hoyt, W.T., & Miller, L. (2015). Mindfulness interventions with youth: A meta-analysis. *Mindfulness*, 6, 290-302. doi: 10.1007/s12671-013-0260-4

# SELF-REPORT MINDFULNESS VALIDITY

Table 1. Within-group responsiveness, by condition and type of outcome

Time Period	Outcome Type	Condition	N	k	ES	Meta-analysis			p	$k_{imp}$	Sensitivity analysis	
						95% CI	$I^2$	Q			ES <sub>adj</sub>	95% CI
Pre-post	Mindfulness	Mindfulness	2134	52	0.49	[0.39, 0.58]	76.87	202.21	< .001	11	0.37	[0.25, 0.49]
	Mindfulness	Waitlist	997	29	-0.03 <sub>a</sub>	[-0.09, 0.03]	0.00	27.53	.490	4	-0.003	[-0.06, 0.05]
	Mindfulness	<i>Bona fide</i>	1479	31	0.23 <sub>a</sub>	[0.10, 0.37]	82.14	115.53	< .001	0		
Pre-FU	Mindfulness	Mindfulness	832	17	0.31	[0.17, 0.45]	72.54	51.77	< .001	0		
	Mindfulness	Waitlist	301	8	-0.06 <sub>b</sub>	[-0.31, 0.20]	77.25	23.51	.001	0		
	Mindfulness	<i>Bona fide</i>	700	13	0.15 <sub>b</sub>	[0.03, 0.27]	54.75	30.47	.002	0		
Pre-post	Clinical Outcomes	Mindfulness	2104	51	0.61	[0.50, 0.72]	83.91	233.66	< .001	0		
	Clinical Outcomes	Waitlist	1011	30	0.12 <sub>c</sub>	[0.06, 0.18]	15.32	30.50	.389	0		
	Clinical Outcomes	<i>Bona fide</i>	1439	29	0.58 <sub>c</sub>	[0.35, 0.81]	95.07	231.03	< .001	0		
Pre-FU	Clinical Outcomes	Mindfulness	939	21	0.65	[0.50, 0.81]	81.85	93.87	< .001	3	0.58	[0.40, 0.75]
	Clinical Outcomes	Waitlist	341	10	0.26 <sub>d</sub>	[0.10, 0.42]	62.62	24.44	.004	2	0.32	[0.16, 0.49]
	Clinical Outcomes	<i>Bona fide</i>	755	15	0.56 <sub>d</sub>	[0.32, 0.81]	92.00	136.79	< .001	0		

Note: Number of studies ( $k$ ) varied from full sample due to data being unavailable to compute pre-post or pre- to follow-up effect sizes for the mindfulness and comparison groups separately. FU = follow-up; ES = effect size (Hedges'  $g_{within}$  quantifies within-group change from baseline); CI = confidence interval;  $I^2$  = heterogeneity;  $Q$  =  $Q$ -statistic assessing degree of heterogeneity;  $p$  =  $p$ -value for  $Q$ -statistic;  $k_{imp}$  = number of studies imputed using trim-and-fill analyses to account for funnel plot asymmetry; ES<sub>adj</sub> = trim-and-fill adjusted effect size; <sub>a</sub> = significant difference between comparison group effect sizes ( $Q[1] = 16.50, p < .001$ ); <sub>b</sub> = no significant difference between comparison group effect sizes ( $Q[1] = 1.68, p = .195$ ); <sub>c</sub> = significant difference between comparison group effect sizes ( $Q[1] = 17.66, p < .001$ ); <sub>d</sub> = no significant difference between comparison group effect sizes ( $Q[1] = 3.45, p = .063$ ). Type of mindfulness measure (FFMQ/KIMS, MAAS, or other) was tested as a moderator of within-group effects on mindfulness. There was no evidence for moderation by the measure of mindfulness used ( $ps > .050$ ).

# SELF-REPORT MINDFULNESS VALIDITY

Table 2. Relative responsiveness (mindfulness versus comparison conditions), by outcome type

Time Period	Outcome Type	Comparison	N	k	ES	Meta-analysis		I <sup>2</sup>	Q	p	Sensitivity analysis		
						95% CI					k <sub>imp</sub>	ES <sub>adj</sub>	95% CI
Pre-post	Mindfulness	Waitlist	1415	25	0.52 <sub>a</sub>	[0.40, 0.64]		17.45	30.40	.172	6	0.45	[0.33, 0.57]
	Mindfulness	<i>Bona fide</i>	2863	30	0.25 <sub>a</sub>	[0.11, 0.38]		59.85	84.59	< .001	0		
Pre-FU	Mindfulness	Waitlist	234	5	0.52 <sub>b</sub>	[0.20, 0.84]		21.65	5.44	.245	2	0.35	[-0.03, 0.72]
	Mindfulness	<i>Bona fide</i>	1430	13	0.10 <sub>b</sub>	[-0.08, 0.28]		58.54	28.40	.005	5	-0.09	[-0.30, 0.12]
Pre-post	Clinical Outcomes	Waitlist	1415	25	0.37 <sub>c</sub>	[0.25, 0.50]		36.34	41.43	.015	4	0.31	[0.16, 0.45]
	Clinical Outcomes	<i>Bona fide</i>	2863	30	0.16 <sub>c</sub>	[0.07, 0.24]		18.67	63.68	< .001	1	0.17	[0.08, 0.26]
Pre-FU	Clinical Outcomes	Waitlist	296	6	0.27 <sub>d</sub>	[-0.14, 0.67]		70.34	16.87	.005	0		
	Clinical Outcomes	<i>Bona fide</i>	2020	19	0.18 <sub>d</sub>	[0.06, 0.30]		42.71	31.83	.023	0		

Note: The combined sample size ( $n = 4,278$  from pre- to post, ) is smaller than the total unique sample size ( $n = 4,743$ ) due to some studies including multiple comparison group types. When multiple comparison conditions were used, the less rigorous comparison condition (i.e., waitlist) was excluded in between-group analyses. ES = effect size ( $\Delta$  quantifies degree to which responsiveness in the mindfulness condition exceeds that in the comparison condition); FU = follow-up; CI = confidence interval;  $I^2$  = heterogeneity;  $Q$  =  $Q$ -statistic assessing degree of heterogeneity;  $p$  =  $p$ -value for  $Q$ -statistic;  $k_{imp}$  = number of studies imputed using trim-and-fill analyses to account for funnel plot asymmetry; ES<sub>adj</sub> = trim-and-fill adjusted effect size; <sub>a</sub> = significant difference between comparison group effect sizes ( $Q[1] = 9.08, p = .003$ ); <sub>b</sub> = significant difference between comparison group effect sizes ( $Q[1] = 4.86, p = .028$ ); <sub>c</sub> = significant difference between comparison group effect sizes ( $Q[1] = 8.27, p = .004$ ); <sub>d</sub> = no significant difference between comparison group effect sizes ( $Q[1] = 0.02, p = .875$ ).

# SELF-REPORT MINDFULNESS VALIDITY

Table 3. Differential responsiveness of mindfulness and clinical outcomes, by condition

Time Period	Outcome Type	Condition	N	k	ES	Meta-analysis		Q	p	Sensitivity analysis		
						95% CI	I <sup>2</sup>			k <sub>imp</sub>	ES <sub>adj</sub>	95% CI
Pre-post	Mindfulness	Mindfulness	2083	50	-0.11	[-0.22, -0.01]	79.88	248.53	< .001	0		
	Mindfulness	Waitlist	997	29	-0.18 <sub>a</sub>	[-0.25, -0.10]	39.98	45.62	.019	0		
	Mindfulness	<i>Bona fide</i>	1439	29	-0.30 <sub>a</sub>	[-0.43, -0.16]	82.16	127.83	< .001	0		
Pre-FU	Mindfulness	Mindfulness	764	16	-0.35	[-0.58, -0.11]	90.20	111.79	< .001	0		
	Mindfulness	Waitlist	301	8	-0.32 <sub>b</sub>	[-0.65, 0.01]	88.31	27.83	< .001	0		
	Mindfulness	<i>Bona fide</i>	633	12	-0.47 <sub>b</sub>	[-0.84, -0.09]	95.84	117.09	< .001	0		

Note: ES = effect size ( $\Delta g$ ) quantifies degree to which responsiveness of mindfulness measures exceeds that for clinical outcome measures for the specified experimental condition (computed as  $g^{mindfulness} - g^{clinical}$ ); FU = follow-up; CI = confidence interval; I<sup>2</sup> = heterogeneity; Q = Q-statistic assessing degree of heterogeneity; p = p-value for Q-statistic; k<sub>imp</sub> = number of studies imputed using trim-and-fill analyses to account for funnel plot asymmetry; ES<sub>adj</sub> = trim-and-fill adjusted effect size; <sub>a</sub> = no significant difference between comparison group effect sizes (Q[1] = 1.57, p = .210); <sub>b</sub> = no significant difference between comparison group effect sizes (Q[1] = 0.22, p = .640). Type of mindfulness measure (FFMQ/KIMS, MAAS, or other) was tested as a moderator of differential responsiveness. There was no evidence for moderation by the measure of mindfulness used (ps > .050).

**Supplemental Materials**

Supplemental Materials Table 1. List of disorders and recognized evidence-based treatments

APA Large Category	Collapsed Disorder Category	Subcategory	Treatments
ADHD	ADHD		CBT
Bipolar	Bipolar		Psychoed, Systematic Care, Cog Ther, Family-Focused Ther, Interpersonal and Social Rhythm Ther
Borderline PD	Borderline		DBT, Mentalization-Based Therapy, Schema-Focused Ther, Transference-Focused Ther
Child and Adolescent Disorders			NA
Chronic or Persistent Pain	Pain	Fibromyalgia	Multicomponent CBT for FM
Chronic or Persistent Pain	Pain	Chronic low back pain	Behavioral Therapy, CBT for CLBP
Chronic or Persistent Pain	Pain	Rheumatologic pain	Multicomponent CBT for Rheumatic Pain
Chronic or Persistent Pain	Pain	Headache	CBT for Chronic Headache
Chronic or Persistent Pain	Pain	Chronic or persistent pain	ACT for Chronic Pain
Depression	Depression		BT/BA, CT, Cog Behav Analysis System of Psychotherapy, IPT, Problem-Solving Ther, Self-Management / Self-Control Therapy, ACT, Behavioral Couple Ther, EFT, REBT, Reminiscence / Life Review Therapy, Self-System Therapy, Short-Term Psychodynamic Therapy
Eating Disorders and Obesity	Weight/Eating	Anorexia Nervosa	Family-Based Treatment, CBT

## SELF-REPORT MINDFULNESS VALIDITY

Eating Disorders and Obesity	Weight/Eating	Bulimia Nervosa	CBT, IPT, Family-Based Treatment, Healthy-Weight Program
Eating Disorders and Obesity	Weight/Eating	Binge Eating Disorder	CBT, IPT
Eating Disorders and Obesity	Weight/Eating	Obesity and Pediatric Overweight	Behavioral Weight Loss Treatment
Generalized Anxiety Disorder	Anxiety		CBT
Insomnia	Sleep		CBT, Sleep Restriction Therapy, Stimulus Control Therapy, Relaxation Training, Paradoxical Intention, EMG Biofeedback
Mixed Anxiety	Anxiety		ACT for Mixed Anxiety Conditions
OCD	Anxiety		Exposure and Response Prevention, CT, ACT
Panic Disorder	Anxiety		CBT, Applied Relaxation, Psychoanalytic Treatment
PTSD	PTSD		PE, PCT, CPT, Seeking Safety, Stress Inoculation Therapy, EMDR, Psych Debriefing
Schizophrenia and Other Mental Illnesses	Schizophrenia		Social Skills Training, CBT, Assertive Community Treatment, Family Psychoed, Supported Employment, Social learning / Token Economy Programs, Cog Remediation, ACT for Psychosis, Cog Adaptation Training, Illness Management and Recovery
Social Phobia and Public Speaking Anxiety	Anxiety		CBT
Specific Phobias	Anxiety		Exposure therapies
Substance and Alcohol Use Disorders	Addiction	Mixed substance abuse	MI, MET, MET + CBT, Prize-Based Contingency Management, Seeking Safety, Friends Care, Guided Self-Change
Substance and Alcohol Use Disorders	Addiction	Alcohol	Behavioral Couples Therapy for Alcohol Use Disorders, Moderate Drinking, Prize-Based Contingency Management

## SELF-REPORT MINDFULNESS VALIDITY

Substance and Alcohol Use Disorders	Addiction	Cocaine	Prize-Based Contingency Management
Substance and Alcohol Use Disorders	Smoking	Smoking	Smoking Cessation with Weight Gain Prevention
Other society's treatments		Smoking	Freedom From Smoking (American Lung Association), CBT (U.S. Department of Health and Human Services)
		Addiction	CBT (National Institute of Drug Abuse)

Note: APA = American Psychological Association; ADHD = Attention Deficit Hyperactivity Disorder; CBT = Cognitive Behavioral Therapy; Psychoed = Psychoeducation; PD = Personality Disorder; Cog = Cognitive; Ther = Therapy; FM = Fibromyalgia; CLBP = Chronic Low Back Pain; BT = Behavioral Therapy; BA = Behavioral Activation; CT = Cognitive Therapy; EFT = Emotion-Focused Therapy; REBT = Rational Emotive Behavior Therapy; IPT = Interpersonal Psychotherapy; EMG = Electromyography; OCD = Obsessive-Compulsive Disorder; ACT = Acceptance and Commitment Therapy; PTSD = Posttraumatic Stress Disorder; PE = Prolonged Exposure; PCT = Present-Centered Therapy; CPT = Cognitive Processing Therapy; EMDR = Eye Movement Desensitization and Reprocessing; MI = Motivational Interviewing; MET = Motivational Enhancement Therapy; U.S. = United States.

# SELF-REPORT MINDFULNESS VALIDITY

Supplemental Materials Table 2. Included studies demographic information and design characteristics

Original Study	Age	% Fem	% Some College	Country	Diagnosis	Time Match	ITT	Mindfulness
Alberts 2012	48.5	100	NA	Netherlands	Obesity	no	yes	KIMS
Alexander 2012	45.00	72	66	United States	Depression	yes	no	MSE
Atkinson 2016	20.57	100	NA	Australia	Eating Disorder	yes	no	FFMQ
Bedard 2014	46.77	45	61	Canada	Depression	no	no	PMS
Bowen 2009	40.50	36.3	NA	United States	Substance Abuse	yes	no	FFMQ
Brewer 2009	38.20	28	52	United States	Substance Abuse	yes	no	FFMQ
Brown 2013	46.61	75	NA	UK	Chronic Pain	no	yes	MAAS
Cathcart 2014	45.54	62.75	NA	Australia	Headache	no	no	FFMQ
Chadwick 2009	41.60	NA	NA	UK	Schizophrenia	no	no	SMQ, SMVQ
Chavooshi 2016	32.67	30	NA	Iran	Chronic Pain	yes	no	MAAS
Cherkin 2016	49.30	65.7	92.3	United States	Back Pain	yes	yes	FFMQ
Chiesa 2015	48.95	72.09	35	Italy	Depression	yes	yes	FFMQ
Colgan 2016	52.00	5.88	NA	United States	PTSD	yes	no	FFMQ
Daubenmier 2011	40.89	100	NA	United States	Obesity	no	yes	KIMS
Davis 2013	21.93	29.1	NA	United States	Smoking	yes	yes	FMI
Davis 2014	44.69	48	59.9	United States	Smoking	yes	yes	FFMQ
Day 2014	41.70	88.9	NA	United States	Headache	no	yes	MAAS
Eisendrath 2016	46.16	76.3	NA	United States	Depression	yes	yes	FFMQ
Garland 2010	40.30	20.8	NA	United States	Alcohol Abuse	yes	no	FFMQ
Garland 2014	48.34	67.83	70.43	United States	Chronic Pain	yes	yes	FFMQ
Garland 2016	37.56	100	NA	United States	Substance Abuse	yes	yes	FFMQ
Goldin 2016	32.70	55.6	NA	United States	Social Anxiety Disorder	yes	yes	FFMQ
Hanstede 2008	25.70	70.59	100	Netherlands	OCD	no	no	SMQ
Hepark 2015	35.89	54.54	NA	Netherlands	ADHD	no	yes	KIMS
Hoge 2013	39.16	50.56	NA	United States	Generalized Anxiety Disorder	yes	yes	FFMQ
Imani 2015	37.41	3.4	7.5	Iran	Opioid dependence	no	no	FFMQ
Jazaieri 2012	32.87	51.79	NA	United States	Social Anxiety Disorder	yes	yes	KIMS
Kearney 2013	52.00	21.28	NA	United States	PTSD	no	yes	FFMQ



# SELF-REPORT MINDFULNESS VALIDITY

Keune 2011	47.16	74.03	NA	Germany	Depression	no	no	FMI
Kocovski 2013	34.00	54	NA	Canada	Social Anxiety Disorder	yes	yes	FMI
Koszycki 2016	39.77	79.54	NA	Canada	Social Anxiety Disorder	no	yes	FFMQ
Kuyken 2008	49.16	76.5	21.51	UK	Depression	no	yes	KIMS
Langer 2012	34.21	41.26	NA	Spain	Schizophrenia	no	no	SMQ
Lopez-Navarro 2015	38.84	17.2	NA	Spain	Schizophrenia	no	yes	MAAS
Manicavasgar 2011	45.84	64.44	NA	Australia	Depression	yes	no	MAAS
Mann 2016	36.20	95	92.5	UK	Depression	no	no	FFMQ
McIndoo 2016	19.22	62.04	100	United States	Depression	yes	yes	FFMQ
Miller 2014	53.95	63.48	53.87	United States	Obesity	yes	no	FFMQ
Moore 2016	70.95	75.57	NA	United States	Anxiety	yes	no	CAMS-R
Morone 2009	75.29	62.86	77.14	United States	Back Pain	yes	no	FFMQ
Morone 2016	74.5	66.3	75.5	United States	Back Pain	yes	yes	MAAS
Nakamura 2013	51.18	78.95	NA	United States	Sleep Disturbance	yes	yes	FFMQ
Perich 2013	NA	65.49	77.47	Australia	Bipolar	no	yes	MAAS
Polusny 2015	58.5	16	NA	United States	PTSD	no	yes	FFMQ
Possemato 2016	46.4	12.9	NA	United States	PTSD	no	yes	MAAS/FFMQ
Pradhan 2007	54.48	87.3	61.9	United States	Rheumatoid Arthritis	no	yes	MAAS
Schmidt 2011	52.5	100	31.59	Germany	Fibromyalgia	yes	yes	FMI
Segal 2010	44.00	53	NA	Canada	Depression	no	yes	TMS
Shahar 2010	46.6	84.44	NA	United States	Depression / Sleep	no	no	MAAS
Strauss 2012	43.00	71.43	NA	UK	Depression	no	yes	SMQ
VanAalderen 2012	47.50	71	59.47	Netherlands	Depression	no	yes	KIMS
Vollestad 2011	42.50	67.11	NA	Norway	Anxiety	no	yes	FFMQ
Wells 2014	45.57	89.47	94.74	United States	Headache	no	yes	FFMQ
Wong 2016	50.00	79.1	33.7	China	Anxiety	yes	yes	FFMQ
Zgierska 2016	51.80	80	77.1	United States	Chronic Pain	no	yes	MAAS

Note: ITT = intent-to-treat analysis reported; Obj = non-self-report measure included; UK = United Kingdom; PTSD = post-traumatic stress disorder; OCD = obsessive-compulsive disorder; ADHD = attention-deficit hyperactivity disorder; NA = not applicable (i.e., data were unavailable); KIMS = Kentucky Inventory of Mindfulness Skills; MSE = Mindfulness Self-Efficacy Rating; FFMQ = Five

## SELF-REPORT MINDFULNESS VALIDITY

Facet Mindfulness Questionnaire; PMS = Philadelphia Mindfulness Scale; TMS = Toronto Mindfulness Scale; MAAS = Mindful Attention Awareness Scale; SMQ = Southhampton Mindfulness Questionnaire; SMVQ = Southhampton Mindfulness Voices Questionnaire; FMI = Freiburg Mindfulness Inventory; CAMS-R = Cognitive Affective Mindfulness Scale - Revised.

# SELF-REPORT MINDFULNESS VALIDITY

Supplemental Materials Table 3. Study descriptive statistics and effects on self-report measures of mindfulness

Original Study	Cont	N	Pre-Post								Pre- to Follow-up							
			$g_m$	$g_m$ var	$g_c$	$g_c$ var	$g$	var	$\Delta_2$	$\Delta_2$ var	$g_m$	$g_m$ var	$g_c$	$g_c$ var	$g$	var	$\Delta_2$	$\Delta_2$ var
Alberts 2012	Waitlist	26	1.04	0.12	0.18	0.06	0.86	0.18	0.16	0.14	NA	NA	NA	NA	NA	NA	NA	NA
Alexander 2012	Bona Fide	134	2.74	0.14	1.59	0.08	1.15	0.22	0.96	0.30	NA	NA	NA	NA	NA	NA	NA	NA
Atkinson 2016	Waitlist	33	0.67	0.11	-0.02	0.06	0.69	0.16	-0.13	0.13	0.58	0.10	0.08	0.06	0.50	0.16	0.08	0.13
Atkinson 2016	Bona Fide	33	0.67	0.11	-0.22	0.08	0.89	0.19	0.55	0.15	0.58	0.10	0.38	0.08	0.20	0.18	0.14	0.15
Bedard 2014	Waitlist	105	0.28	0.02	0.12	0.02	0.16	0.04	-0.05	0.04	NA	NA	NA	NA	NA	NA	NA	NA
Bowen 2009	Bona Fide	163	0.14	0.02	-0.17	0.02	0.3	0.04	0.03	0.04	0.00	0.02	0.15	0.02	-0.15	0.04	-0.17	0.04
Brewer 2009	Bona Fide	36	0.69	0.12	0.26	0.13	0.43	0.25	0.94	0.23	NA	NA	NA	NA	NA	NA	NA	NA
Brown 2013	Waitlist	40	0.77	0.08	-0.06	0.07	0.83	0.15	0.34	0.12	NA	NA	NA	NA	NA	NA	NA	NA
Cathcart 2014	Waitlist	58	-0.05	0.04	0.03	0.05	-0.09	0.09	-0.1	0.07	NA	NA	NA	NA	NA	NA	NA	NA
Chadwick 2009	Waitlist	22	NA	NA	NA	NA	0.64	0.16	0.36	0.15	NA	NA	NA	NA	NA	NA	NA	NA
Chavooshi 2016	Bona Fide	43	0.75	0.06	3.64	0.36	-2.89	0.42	1.31	0.58	NA	NA	NA	NA	NA	NA	NA	NA
Cherkin 2016	Waitlist	229	0.27	0.01	0.05	0.01	0.22	0.01	-0.13	0.01	0.26	0.01	0.21	0.01	0.05	0.01	-0.25	0.01
Cherkin 2016	Bona Fide	229	0.27	0.01	0.24	0.01	0.03	0.01	-0.05	0.01	0.26	0.01	0.25	0.01	0.01	0.01	0.05	0.01
Chiesa 2015	Bona Fide	50	0.32	0.03	0.04	0.03	0.29	0.05	-0.26	0.07	0.35	0.03	0	0.03	0.35	0.05	-0.46	0.07
Colgan 2016	Bona Fide	56	0.18	0.04	0.11	0.04	0.07	0.08	-0.38	0.08	NA	NA	NA	NA	NA	NA	NA	NA
Daubenmier 2011	Waitlist	47	0.39	0.03	0.00	0.03	0.39	0.05	0.29	0.05	NA	NA	NA	NA	NA	NA	NA	NA
Davis 2013	Bona Fide	55	-0.30	0.06	-0.45	0.08	0.16	0.14	-0.62	0.14	NA	NA	NA	NA	NA	NA	NA	NA
Davis 2014	Bona Fide	135	0.44	0.04	0.47	0.04	-0.04	0.07	-0.37	0.06	0.59	0.04	0.45	0.04	0.13	0.07	-0.1	0.06
Day 2014	Waitlist	36	0.11	0.05	-0.40	0.06	0.50	0.11	0.33	0.09	NA	NA	NA	NA	NA	NA	NA	NA
Eisendrath 2016	Bona Fide	173	0.61	0.01	0.40	0.01	0.21	0.02	0.15	0.02	NA	NA	NA	NA	NA	NA	NA	NA
Garland 2010	Bona Fide	53	0.11	0.05	0.14	0.05	-0.03	0.10	-0.01	0.09	NA	NA	NA	NA	NA	NA	NA	NA
Garland 2014	Bona Fide	115	0.51	0.02	0.00	0.02	0.50	0.04	0.05	0.03	NA	NA	NA	NA	NA	NA	NA	NA
Garland 2016	Bona Fide	128	0.53	0.02	-0.01	0.02	0.54	0.03	0.21	0.03	NA	NA	NA	NA	NA	NA	NA	NA
Garland 2016	Bona Fide	128	0.53	0.02	0.09	0.02	0.43	0.04	0.21	0.04	NA	NA	NA	NA	NA	NA	NA	NA
Goldin 2016	Waitlist	72	0.91	0.04	-0.35	0.03	1.26	0.07	0.34	0.06	NA	NA	NA	NA	NA	NA	NA	NA
Goldin 2016	Bona Fide	72	0.91	0.04	0.87	0.04	0.04	0.08	0.64	0.09	NA	NA	NA	NA	NA	NA	NA	NA
Hanstede 2008	Waitlist	17	1.42	0.22	-0.19	0.09	1.61	0.32	0.06	0.34	NA	NA	NA	NA	NA	NA	NA	NA

# SELF-REPORT MINDFULNESS VALIDITY

Hepark 2015	Waitlist	103	0.68	0.02	-0.03	0.02	0.71	0.04	0.23	0.04	NA	NA	NA	NA	NA	NA	NA	NA
Hoge 2013	Bona Fide	93	0.83	0.07	0.34	0.05	0.48	0.12	0.15	0.09	NA	NA	NA	NA	NA	NA	NA	NA
Imani 2015	Waitlist	30	NA	NA	NA	NA	0.67	0.09	0.18	0.09	NA	NA	NA	NA	NA	NA	NA	NA
Jazaieri 2012	Bona Fide	56	1.16	0.07	0.43	0.06	0.73	0.12	0.41	0.10	NA	NA	NA	NA	NA	NA	NA	NA
Kearney 2013	Waitlist	47	0.48	0.04	-0.06	0.04	0.54	0.09	0.33	0.09	0.38	0.04	-0.17	0.04	0.55	0.09	0.37	0.09
Keune 2011	Waitlist	91	1.00	0.04	0.23	0.03	0.77	0.06	0.26	0.05	NA	NA	NA	NA	NA	NA	NA	NA
Kocovski 2013	Waitlist	106	0.57	0.02	0.00	0.03	0.57	0.05	-0.09	0.05	NA	NA	NA	NA	NA	NA	NA	NA
Kocovski 2013	Bona Fide	106	0.57	0.02	0.30	0.02	0.27	0.04	0.35	0.05	NA	NA	NA	NA	NA	NA	NA	NA
Koszycki 2016	Waitlist	39	0.43	0.05	-0.35	0.05	0.78	0.10	-0.39	0.11	NA	NA	NA	NA	NA	NA	NA	NA
Kuyken 2008	Bona Fide	123	0.87	0.03	0.18	0.02	0.68	0.04	0.41	0.04	NA	NA	NA	NA	NA	NA	NA	NA
Langer 2012	Waitlist	23	1.14	0.20	0.19	0.08	0.95	0.28	0.24	0.27	NA	NA	NA	NA	NA	NA	NA	NA
Lopez-Navarro 2015	Waitlist	44	0.18	0.04	-0.15	0.04	0.33	0.09	0.06	0.07	NA	NA	NA	NA	NA	NA	NA	NA
Manicavasgar 2011	Bona Fide	69	0.50	0.05	0.67	0.04	-0.17	0.10	-0.16	0.09	NA	NA	NA	NA	NA	NA	NA	NA
Mann 2016	Waitlist	38	0.79	0.07	0.23	0.08	0.56	0.15	0.36	0.14	1.34	0.11	0.13	0.06	1.21	0.16	0.03	0.15
McIndoo 2016	Waitlist	36	0.68	0.06	-0.01	0.06	0.69	0.12	-0.16	0.11	-0.19	0.05	-1.25	0.12	1.06	0.17	0.43	0.14
McIndoo 2016	Bona Fide	36	0.68	0.06	0.25	0.06	0.44	0.12	0.67	0.12	-0.19	0.05	-1.13	0.1	0.94	0.14	1.53	0.14
Miller 2014	Bona Fide	68	0.32	0.04	0.16	0.04	0.16	0.07	0.33	0.06	0.49	0.04	0.28	0.04	0.21	0.08	0.32	0.06
Moore 2016	Bona Fide	67	0.65	0.03	0.24	0.02	0.40	0.05	0.17	0.05	NA	NA	NA	NA	NA	NA	NA	NA
Morone 2009	Bona Fide	40	-0.06	0.03	0.07	0.03	-0.13	0.06	-0.28	0.07	-0.06	0.03	0.19	0.03	-0.25	0.06	-0.43	0.07
Morone 2016	Bona Fide	282	-0.14	0.01	0.14	0.01	-0.28	0.01	-0.54	0.01	0	0.01	0.28	0.01	-0.28	0.01	-0.45	0.01
Nakamura 2013	Bona Fide	38	0.23	0.05	0.03	0.05	0.2	0.1	-0.61	0.10	0.49	0.05	0.34	0.05	0.14	0.11	-0.84	0.15
Perich 2013	Waitlist	95	0.59	0.02	0.16	0.02	0.43	0.04	0.49	0.04	0.55	0.02	0.26	0.02	0.29	0.04	0.68	0.04
Polusny 2015	Bona Fide	116	0.59	0.02	0.06	0.02	0.52	0.04	0.25	0.03	0.57	0.02	0.03	0.02	0.54	0.04	0.09	0.03
Possemato 2016	Waitlist	62	0.17	0.03	-0.02	0.02	0.19	0.05	0.08	0.05	NA	NA	NA	NA	NA	NA	NA	NA
Pradhan 2007	Waitlist	63	0.01	0.03	0.13	0.03	-0.12	0.06	0.10	0.07	NA	NA	NA	NA	NA	NA	NA	NA
Schmidt 2011	Bona Fide	118	0.50	0.02	0.15	0.02	0.35	0.04	0.30	0.03	0.22	0.02	-0.09	0.02	0.32	0.04	0.08	0.03
Schmidt 2011	Waitlist	118	0.50	0.02	-0.11	0.02	0.61	0.04	0.42	0.03	0.22	0.02	-0.08	0.02	0.30	0.04	0.18	0.03
Segal 2010	Bona Fide	54	0.65	0.05	-0.51	0.05	1.16	0.10	1.06	0.11	NA	NA	NA	NA	NA	NA	NA	NA
Shahar 2010	Waitlist	52	0.43	0.04	-0.59	0.06	1.02	0.10	0.67	0.09	NA	NA	NA	NA	NA	NA	NA	NA
Strauss 2012	Waitlist	28	0.73	0.08	-0.40	0.07	1.13	0.15	-0.37	0.17	NA	NA	NA	NA	NA	NA	NA	NA
VanAalderen 2012	Waitlist	219	0.51	0.01	-0.04	0.01	0.54	0.01	0.00	0.01	NA	NA	NA	NA	NA	NA	NA	NA

## SELF-REPORT MINDFULNESS VALIDITY

Vollestad 2011	Waitlist	76	0.52	0.03	-0.09	0.03	0.61	0.05	0.08	0.05	NA	NA	NA	NA	NA	NA	NA	NA
Wells 2014	Waitlist	19	0.34	0.09	-0.32	0.10	0.66	0.19	-0.04	0.16	0.57	0.10	-0.26	0.09	0.83	0.19	0.29	0.16
Wong 2016	Bona Fide	122	0.16	0.02	0.21	0.02	-0.05	0.03	0.02	0.03	0.03	0.02	0.21	0.02	-0.18	0.03	-0.18	0.03
Wong 2016	Waitlist	122	0.16	0.02	-0.10	0.02	0.26	0.03	-0.05	0.03	NA	NA	NA	NA	NA	NA	NA	NA
Zgierska 2016	Waitlist	35	NA	NA	NA	NA	0.44	0.12	-0.02	0.1	NA	NA	NA	NA	0.19	0.13	-0.27	0.10

Note: Cont = control condition type; N = combined mindfulness and comparison group intent-to-treat sample size;  $g_m$  = within-group effect size for mindfulness conditions;  $g_c$  = within-group effects for comparison conditions; var = variance; between-group effects are represented by  $g$  and var without subscripts;  $\Delta_2$  = relative responsiveness of mindfulness measures and clinical outcomes; NA = not applicable (i.e., data were unavailable to compute the given effect); *Bona fide* = comparison group was intended to be therapeutic.

# SELF-REPORT MINDFULNESS VALIDITY

Supplemental Materials Table 4. Study descriptive statistics and effects on clinical outcomes

Original Study	Cont	N	Pre-Post						Pre- to Follow-up					
			$g_m$	$g_m$ var	$g_c$	$g_c$ var	$g$	var	$g_m$	$g_m$ var	$g_c$	$g_c$ var	$g$	var
Alberts 2012	Waitlist	26	0.67	0.06	-0.02	0.04	0.69	0.09	NA	NA	NA	NA	NA	NA
Alexander 2012	Bona Fide	134	3.00	0.17	2.81	0.18	0.19	0.35	NA	NA	NA	NA	NA	NA
Atkinson 2016	Waitlist	33	0.79	0.06	-0.03	0.03	0.82	0.09	0.54	0.05	0.12	0.03	0.41	0.09
Atkinson 2016	Bona Fide	33	0.79	0.06	0.45	0.05	0.34	0.11	0.54	0.05	0.48	0.05	0.06	0.1
Bedard 2014	Waitlist	105	0.41	0.02	0.20	0.02	0.21	0.04	NA	NA	NA	NA	NA	NA
Bowen 2009	Bona Fide	163	1.17	0.02	0.90	0.02	0.27	0.04	0.98	0.01	0.95	0.02	0.03	0.03
Brewer 2009	Bona Fide	36	NA	NA	NA	NA	-0.51	0.2	NA	NA	NA	NA	NA	NA
Brown 2013	Waitlist	40	0.47	0.04	-0.02	0.04	0.50	0.08	NA	NA	NA	NA	NA	NA
Cathcart 2014	Waitlist	58	0.17	0.02	0.16	0.03	0.01	0.05	NA	NA	NA	NA	NA	NA
Chadwick 2009	Waitlist	22	NA	NA	NA	NA	0.27	0.13	NA	NA	NA	NA	NA	NA
Chavooshi 2016	Bona Fide	43	0.77	0.06	4.97	0.63	-4.2	0.7	NA	NA	NA	NA	NA	NA
Cherkin 2016	Waitlist	229	0.61	0.01	0.26	0.01	0.32	0.01	0.78	0.01	0.48	0.01	0.30	0.01
Cherkin 2016	Bona Fide	229	0.61	0.01	0.53	0.01	0.04	0.01	0.78	0.01	0.82	0.01	-0.04	0.01
Chiesa 2015	Bona Fide	50	0.87	0.04	0.33	0.04	0.54	0.08	1.09	0.05	0.29	0.04	0.81	0.09
Colgan 2016	Bona Fide	56	0.57	0.04	0.11	0.04	0.46	0.08	NA	NA	NA	NA	NA	NA
Daubenmier 2011	Waitlist	47	0.08	0.02	-0.01	0.02	0.09	0.05	NA	NA	NA	NA	NA	NA
Davis 2013	Bona Fide	55	NA	NA	NA	NA	0.78	0.15	NA	NA	NA	NA	NA	NA
Davis 2014	Bona Fide	135	0.48	0.02	0.15	0.02	0.18	0.03	NA	NA	NA	NA	0.23	0.05
Day 2014	Waitlist	36	0.26	0.03	0.08	0.03	0.17	0.06	NA	NA	NA	NA	NA	NA
Eisendrath 2016	Bona Fide	173	0.85	0.02	0.79	0.02	0.29	0.03	NA	NA	NA	NA	0.05	0.02
Garland 2010	Bona Fide	53	0.27	0.04	0.30	0.04	-0.02	0.08	NA	NA	NA	NA	NA	NA
Garland 2014	Bona Fide	115	0.46	0.01	0.00	0.01	0.46	0.02	0.44	0.01	-0.24	0.01	0.69	0.03
Garland 2016	Bona Fide	128	0.36	0.02	0.03	0.02	0.32	0.03	NA	NA	NA	NA	NA	NA
Garland 2016	Bona Fide	128	0.36	0.02	0.14	0.02	0.22	0.04	NA	NA	NA	NA	NA	NA
Goldin 2016	Waitlist	72	1.23	0.04	0.32	0.02	0.91	0.06	NA	NA	NA	NA	NA	NA
Goldin 2016	Bona Fide	72	1.23	0.04	1.84	0.06	-0.60	0.10	NA	NA	NA	NA	NA	NA
Hanstede 2008	Waitlist	17	1.64	0.27	0.09	0.09	1.55	0.36	NA	NA	NA	NA	NA	NA

# SELF-REPORT MINDFULNESS VALIDITY

Hepark 2015	Waitlist	103	0.62	0.02	0.13	0.02	0.48	0.03	NA	NA	NA	NA	NA	NA
Hoge 2013	Bona Fide	93	1.00	0.02	0.66	0.02	0.38	0.04	NA	NA	NA	NA	NA	NA
Imani 2015	Waitlist	30	NA	NA	NA	NA	0.49	0.1	NA	NA	NA	NA	NA	NA
Jazaieri 2012	Bona Fide	56	1.18	0.04	0.86	0.04	0.32	0.08	1.49	0.09	1.47	0.11	0.02	0.2
Kearney 2013	Waitlist	47	0.60	0.04	0.39	0.05	0.21	0.09	0.4	0.04	0.22	0.05	0.18	0.09
Keune 2011	Waitlist	91	0.59	0.02	0.08	0.02	0.51	0.04	NA	NA	NA	NA	NA	NA
Kocovski 2013	Waitlist	106	0.90	0.02	0.23	0.02	0.66	0.05	NA	NA	NA	NA	NA	NA
Kocovski 2013	Bona Fide	106	0.90	0.02	0.98	0.03	-0.08	0.05	NA	NA	NA	NA	NA	NA
Koszycki 2016	Waitlist	39	1.54	0.07	0.38	0.04	1.16	0.11	NA	NA	NA	NA	NA	NA
Kuyken 2008	Bona Fide	123	0.22	0.01	-0.05	0.01	0.13	0.05	NA	NA	NA	NA	0.27	0.04
Langer 2012	Waitlist	23	1.01	0.18	0.31	0.08	0.71	0.26	NA	NA	NA	NA	NA	NA
Lopez-Navarro 2015	Waitlist	44	0.29	0.03	0.02	0.03	0.27	0.06	NA	NA	NA	NA	NA	NA
Manicavasgar 2011	Bona Fide	69	0.75	0.05	0.76	0.04	-0.01	0.08	1.28	0.18	1.23	0.12	0.05	0.3
Mann 2016	Waitlist	38	0.49	0.06	0.28	0.08	0.21	0.14	1.00	0.08	-0.18	0.06	1.18	0.14
McIndoo 2016	Waitlist	36	1.10	0.05	0.25	0.04	0.85	0.1	1.00	0.05	0.37	0.05	0.62	0.1
McIndoo 2016	Bona Fide	36	1.10	0.05	1.34	0.08	-0.24	0.13	1.00	0.05	1.58	0.09	-0.59	0.14
Miller 2014	Bona Fide	68	0.34	0.02	0.5	0.02	-0.17	0.05	0.38	0.02	0.49	0.02	-0.11	0.05
Moore 2016	Bona Fide	67	0.38	0.02	0.15	0.02	0.23	0.04	NA	NA	NA	NA	NA	NA
Morone 2009	Bona Fide	40	0.4	0.04	0.26	0.03	0.14	0.07	0.43	0.04	0.25	0.03	0.18	0.07
Morone 2016	Bona Fide	282	0.46	0.00	0.20	0.00	0.26	0.01	0.37	0.00	0.20	0.00	0.17	0.01
Nakamura 2013	Bona Fide	38	1.24	0.06	0.43	0.04	0.81	0.11	1.71	0.12	0.73	0.07	0.98	0.19
Perich 2013	Waitlist	95	0.23	0.01	0.30	0.01	-0.09	0.04	0.19	0.01	0.57	0.01	-0.28	0.03
Polusny 2015	Bona Fide	116	0.67	0.02	0.39	0.01	0.17	0.03	0.87	0.02	0.41	0.01	0.35	0.03
Possemato 2016	Waitlist	62	0.52	0.02	0.41	0.03	0.11	0.05	0.43	0.03	0.57	0.04	-0.13	0.07
Pradhan 2007	Waitlist	63	-0.2	0.04	0.02	0.03	-0.22	0.07	NA	NA	NA	NA	NA	NA
Schmidt 2011	Bona Fide	118	0.23	0.01	0.19	0.01	0.05	0.02	0.24	0.01	0.00	0.01	0.24	0.02
Schmidt 2011	Waitlist	118	0.23	0.01	0.05	0.01	0.18	0.02	0.24	0.01	0.12	0.01	0.12	0.02
Segal 2010	Bona Fide	54	0.08	0.06	-0.02	0.05	-0.03	0.08	NA	NA	NA	NA	0.18	0.09
Shahar 2010	Waitlist	52	0.40	0.04	0.05	0.04	0.35	0.08	NA	NA	NA	NA	NA	NA
Strauss 2012	Waitlist	28	1.21	0.12	-0.3	0.07	1.50	0.18	NA	NA	NA	NA	NA	NA
VanAalderen 2012	Waitlist	219	0.51	0.01	-0.03	0.01	0.54	0.01	NA	NA	NA	NA	NA	NA

## SELF-REPORT MINDFULNESS VALIDITY

Vollestad 2011	Waitlist	76	0.58	0.02	0.06	0.02	0.53	0.03	NA	NA	NA	NA	NA	NA
Wells 2014	Waitlist	19	0.57	0.06	-0.13	0.06	0.70	0.12	0.55	0.06	0.01	0.06	0.54	0.12
Wong 2016	Bona Fide	122	0.33	0.01	0.40	0.01	-0.07	0.03	0.55	0.01	0.56	0.01	0.00	0.03
Wong 2016	Waitlist	122	0.33	0.01	0.02	0.01	0.31	0.03	NA	NA	NA	NA	NA	NA
Zgierska 2016	Waitlist	35	0.52	0.04	-0.26	0.08	0.46	0.08	0.37	0.05	-0.05	0.05	0.45	0.07

Note: Cont = control condition type; N = combined mindfulness and comparison group intent-to-treat sample size;  $g_m$  = within-group effect size for mindfulness conditions;  $g_c$  = within-group effects for comparison conditions; var = variance; between-group effects are represented by  $g$  and var without subscripts; NA = not applicable (i.e., data were unavailable to compute the given effect); *Bona fide* = comparison group was intended to be therapeutic.



## SELF-REPORT MINDFULNESS VALIDITY

### Supplemental Materials Table 5. Included Studies

Alberts, H. J. E. M., Thewissen, R., & Raes, L. (2012). Dealing with problematic eating behaviour. The effects of a mindfulness-based intervention on eating behaviour, food cravings, dichotomous thinking and body image concern. *Appetite*, 58(3), 847-851.

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

Atkinson, M. J., & Wade, T. D. (2016). Does mindfulness have potential in eating disorders prevention? A preliminary controlled trial with young adult women. *Early Intervention in Psychiatry*, 10, 234-245.

Bédard, M., Felteau, M., Marshall, S., Cullen, N., Gibbons, C., Dubois, S., ... & Gainer, R. (2014). Mindfulness-based cognitive therapy reduces symptoms of depression in people with a traumatic brain injury: results from a randomized controlled trial. *The Journal of Head Trauma Rehabilitation*, 29(4), E13-E22.

Bieling, P. J., Hawley, L. L., Bloch, R. T., Corcoran, K. M., Levitan, R. D., Young, L. T., ... & Segal, Z. V. (2012). Treatment-specific changes in decentering following mindfulness-based cognitive therapy versus antidepressant medication or placebo for prevention of depressive relapse. *Journal of Consulting and Clinical Psychology*, 80(3), 365-372.

Bowen, S., Chawla, N., Collins, S., Witkiewitz, K., Hsu, S., Grow, J.,...Marlatt, A. (2009). Mindfulness-based relapse prevention for substance use disorders: A pilot efficacy trial. *Substance Abuse*, 30, 295-305.

Brewer, J., Sinha, R., Chen, J., Michalsen, R., Babuscio, T., Nich, C,... Rounsaville, B.J. (2009). Mindfulness training and stress reactivity in substance abuse: Results from a randomized, controlled Stage I pilot study. *Substance Abuse*, 30, 306-317.

Britton, W. B., Haynes, P. L., Fridel, K. W., & Bootzin, R. R. (2010). Polysomnographic and subjective profiles of sleep continuity before and after mindfulness-based cognitive therapy in partially remitted depression. *Psychosomatic Medicine*, 72(6), 539-548.

Britton, W. B., Haynes, P. L., Fridel, K. W., & Bootzin, R. R. (2012). Mindfulness-based cognitive therapy improves polysomnographic and subjective sleep profiles in antidepressant users with sleep complaints. *Psychotherapy and Psychosomatics*, 81(5), 296-304.

Brown, C. A., & Jones, A. K. (2013). Psychobiological correlates of improved mental health in patients with musculoskeletal pain after a mindfulness-based pain management program. *The Clinical Journal of Pain*, 29(3), 233-244.

## SELF-REPORT MINDFULNESS VALIDITY

Cathcart, S., Barone, V., Immink, M., & Proeve, M. (2013). Mindfulness training does not reduce generalized hyperalgesia in chronic tension-type headache. *Journal of Pain Management*, 6(3), 217-221.

Cathcart, S., Galatis, N., Immink, M., Proeve, M., & Petkov, J. (2014). Brief mindfulness-based therapy for chronic tension-type headache: a randomized controlled pilot study. *Behavioural and Cognitive Psychotherapy*, 42, 1-15.

Chadwick, P., Hughes, S., Russell, D., Russell, I., & Dagnan, D. (2009). Mindfulness groups for distressing voices and paranoia: a replication and randomized feasibility trial. *Behavioural and Cognitive Psychotherapy*, 37, 403-412.

Chavooshi, B., Mohammadkhani, P., & Dolatshahee, B. (2016). Efficacy of intensive short-term dynamic psychotherapy for medically unexplained pain: A pilot three-armed randomized controlled trial comparison with mindfulness-based stress reduction. *Psychotherapy and Psychosomatics*, 85(2), 123-125.

Cherkin, D. C., Sherman, K. J., Balderson, B. H., Cook, A. J., Anderson, M. L., Hawkes, R. J., ... & Turner, J. A. (2016). Effect of mindfulness-based stress reduction vs cognitive behavioral therapy or usual care on back pain and functional limitations in adults with chronic low back pain: A randomized clinical trial. *JAMA*, 315(12), 1240-1249.

Chiesa, A., Castagner, V., Andrisano, C., Serretti, A., Mandelli, L., Porcelli, S., & Giommi, F. (2015). Mindfulness-based cognitive therapy vs. psycho-education for patients with major depression who did not achieve remission following antidepressant treatment. *Psychiatry Research*, 226(2), 474-483.

Colgan, D. D., Christopher, M., Michael, P., & Wahbeh, H. (2016). The Body Scan and Mindful Breathing Among Veterans with PTSD: Type of Intervention Moderates the Relationship Between Changes in Mindfulness and Post-treatment Depression. *Mindfulness*, 7(2), 372-383.

Daubenmier, J., Kristeller, J., Hecht, F. M., Maninger, N., Kuwata, M., Jhaveri, K., ... & Epel, E. (2011). Mindfulness intervention for stress eating to reduce cortisol and abdominal fat among overweight and obese women: an exploratory randomized controlled study. *Journal of Obesity*, 2011, 1-13.

Davis, J. M., Mills, D. M., Stankevitz, K. A., Manley, A. R., Majeskie, M. R., & Smith, S. S. (2013). Pilot randomized trial on mindfulness training for smokers in young adult binge drinkers. *BMC Complementary and Alternative Medicine*, 13(1), 1-10.

Davis, J.M., Manley, A.R., Goldberg, S.B., Smith, S.S., & Jorenby, D.E. (2014). Randomized trial comparing mindfulness training for smokers to a matched control. *Journal of Substance Abuse Treatment*, 47(3), 213-221.

## SELF-REPORT MINDFULNESS VALIDITY

Day, M. A., Thorn, B. E., Ward, L. C., Rubin, N., Hickman, S. D., Scogin, F., & Kilgo, G. R. (2014). Mindfulness-based cognitive therapy for the treatment of headache pain: a pilot study. *The Clinical Journal of Pain*, 30(2), 152-161.

Eisendrath, S. J., Gillung, E., Delucchi, K. L., Segal, Z. V., Nelson, J. C., McInnes, L. A., ... & Feldman, M. D. (2016). A randomized controlled trial of mindfulness-based cognitive therapy for treatment-resistant depression. *Psychotherapy and Psychosomatics*, 85(2), 99-110.

Garland, E. L., Gaylord, S. A., Boettiger, C. A., & Howard, M. O. (2010). Mindfulness training modifies cognitive, affective, and physiological mechanisms implicated in alcohol dependence: results of a randomized controlled pilot trial. *Journal of Psychoactive Drugs*, 42(2), 177-192.

Garland, E. L., & Howard, M. O. (2013). Mindfulness-oriented recovery enhancement reduces pain attentional bias in chronic pain patients. *Psychotherapy and Psychosomatics*, 82(5), 311-318.

Garland, E.L., Manusov, E.G., Froeliger, B., Kelly, A., Williams, J.M., & Howard, M.O. (2014). Mindfulness-oriented recovery enhancement for chronic pain and prescription opioid misuse: Results from an early-stage randomized controlled trial. *Journal of Consulting and Clinical Psychology*, 82(3), 448-459.

Garland, E. L., Thomas, E., & Howard, M. O. (2014). Mindfulness-oriented recovery enhancement ameliorates the impact of pain on self-reported psychological and physical function among opioid-using chronic pain patients. *Journal of Pain and Symptom Management*, 48(6), 1091-1099.

Garland, E. L., Roberts-Lewis, A., Tronnier, C. D., Graves, R., & Kelley, K. (2016). Mindfulness-Oriented Recovery Enhancement versus CBT for co-occurring substance dependence, traumatic stress, and psychiatric disorders: Proximal outcomes from a pragmatic randomized trial. *Behaviour Research and Therapy*, 77, 7-16.

Goldin, P., Ziv, M., Jazaieri, H., & Gross, J. (2012). Randomized controlled trial of mindfulness-based stress reduction versus aerobic exercise: effects on the self-referential brain network in social anxiety disorder. *Frontiers in Human Neuroscience*, 6, 1-16.

Goldin, P., Ziv, M., Jazaieri, H., Hahn, K., & Gross, J. J. (2013). MBSR vs aerobic exercise in social anxiety: fMRI of emotion regulation of negative self-beliefs. *Social Cognitive and Affective Neuroscience*, 8(1), 65-72.

Goldin, P. R., Morrison, A., Jazaieri, H., Brozovich, F., Heimberg, R., & Gross, J. J. (2016). Group CBT versus MBSR for social anxiety disorder: A randomized controlled trial. *Journal of Consulting and Clinical Psychology*, 84(5), 427-437.

Grossman, P., Deuring, G., Walach, H., Schwarzer, B., & Schmidt, S. (2017). Mindfulness-based intervention does not influence cardiac autonomic control or the pattern of physical activity in

## SELF-REPORT MINDFULNESS VALIDITY

fibromyalgia during daily life: an ambulatory, multimeasure randomized controlled trial. *The Clinical Journal of Pain*, 33, 385-94.

Hanstede, M., Gidron, Y., & Nyklícek, I. (2008). The effects of a mindfulness intervention on obsessive-compulsive symptoms in a non-clinical student population. *The Journal of Nervous and Mental Disease*, 196(10), 776-779.

Hepark, S., Janssen, L., de Vries, A., Schoenberg, P. L., Donders, R., Kan, C. C., & Speckens, A. E. (2015). The efficacy of adapted MBCT on core symptoms and executive functioning in adults with ADHD: A preliminary randomized controlled trial. *Journal of Attention Disorders*, doi: 10.1177/1087054715613587

Hoge, E.A., Bui, E., Marques, L., Metcalf, C.A., Morris, L.K., Robinaugh, D.J.,...& Simon, N.M. (2013). Randomized controlled trial of mindfulness meditation for generalized anxiety disorder: Effects on anxiety and stress reactivity. *Journal of Clinical Psychiatry*, 74, 786-792.

Hoge, E. A., Bui, E., Goetter, E., Robinaugh, D. J., Ojserkis, R. A., Fresco, D. M., & Simon, N. M. (2015). Change in decentering mediates improvement in anxiety in mindfulness-based stress reduction for generalized anxiety disorder. *Cognitive Therapy and Research*, 39(2), 228-235.

Imani, S., Vahid, M. K. A., Gharraee, B., Noroozi, A., Habibi, M., & Bowen, S. (2015). Effectiveness of Mindfulness-Based Group Therapy Compared to the Usual Opioid Dependence Treatment. *Iranian Journal of Psychiatry*, 10(3), 175.

Jazaieri, H., Goldin, P. R., Werner, K., Ziv, M., & Gross, J. J. (2012). A randomized trial of MBSR versus aerobic exercise for social anxiety disorder. *Journal of Clinical Psychology*, 68(7), 715-731.

Jazaieri, H., Lee, I. A., Goldin, P. R., & Gross, J. J. (2016). Pre-treatment social anxiety severity moderates the impact of mindfulness-based stress reduction and aerobic exercise. *Psychology and Psychotherapy: Theory, Research and Practice*, 89, 229-234.

Kearney, D.J., McDermott, K., Malte, C., Martinez, M., & Simpson, T.L. (2013). Effects of participation in a mindfulness program for veterans with posttraumatic stress disorder: A randomized controlled pilot study. *Journal of Clinical Psychology*, 69, 14-27.

Keune, P. M., Bostanov, V., Hautzinger, M., & Kotchoubey, B. (2011). Mindfulness-based cognitive therapy (MBCT), cognitive style, and the temporal dynamics of frontal EEG alpha asymmetry in recurrently depressed patients. *Biological Psychology*, 88(2), 243-252.

Kocovski, N. L., Fleming, J. E., Hawley, L. L., Huta, V., & Antony, M. M. (2013). Mindfulness and acceptance-based group therapy versus traditional cognitive behavioral group therapy for social anxiety disorder: A randomized controlled trial. *Behaviour Research and Therapy*, 51(12), 889-898.

## SELF-REPORT MINDFULNESS VALIDITY

- Koszycki, D., Thake, J., Mavounza, C., Daoust, J. P., Taljaard, M., & Bradwejn, J. (2016). Preliminary Investigation of a Mindfulness-Based Intervention for Social Anxiety Disorder That Integrates Compassion Meditation and Mindful Exposure. *The Journal of Alternative and Complementary Medicine*, 22(5), 363-374.
- Kuyken, W., Byford, S., Taylor, R. S., Watkins, E., Holden, E., White, K., ... & Teasdale, J. D. (2008). Mindfulness-based cognitive therapy to prevent relapse in recurrent depression. *Journal of Consulting and Clinical Psychology*, 76(6), 966-978.
- Kuyken, W., Watkins, E., Holden, E., White, K., Taylor, R. S., Byford, S., ... & Dalgleish, T. (2010). How does mindfulness-based cognitive therapy work?. *Behaviour Research and Therapy*, 48(11), 1105-1112.
- Langer, Á. I., Cangas, A. J., Salcedo, E., & Fuentes, B. (2012). Applying mindfulness therapy in a group of psychotic individuals: a controlled study. *Behavioural and Cognitive Psychotherapy*, 40, 105-109.
- López-Navarro, E., Del Canto, C., Belber, M., Mayol, A., Fernández-Alonso, O., Lluís, J., ... & Chadwick, P. (2015). Mindfulness improves psychological quality of life in community-based patients with severe mental health problems: A pilot randomized clinical trial. *Schizophrenia Research*, 168(1), 530-536.
- Manicavasagar, V., Parker, G., & Perich, T. (2011). Mindfulness-based cognitive therapy vs cognitive behaviour therapy as a treatment for non-melancholic depression. *Journal of Affective Disorders*, 130(1), 138-144.
- Manicavasagar, V., Perich, T., & Parker, G. (2012). Cognitive predictors of change in cognitive behaviour therapy and mindfulness-based cognitive therapy for depression. *Behavioural and Cognitive Psychotherapy*, 40, 227-232.
- Mann, J., Kuyken, W., O'Mahen, H., Ukoumunne, O. C., Evans, A., & Ford, T. (2016). Manual development and pilot randomised controlled trial of mindfulness-based cognitive therapy versus usual care for parents with a history of depression. *Mindfulness*, 7(5), 1024-1033.
- McIndoo, C. C., File, A. A., Preddy, T., Clark, C. G., & Hopko, D. R. (2016). Mindfulness-based therapy and behavioral activation: A randomized controlled trial with depressed college students. *Behaviour Research and Therapy*, 77, 118-128.
- Miller, C. K., Kristeller, J. L., Headings, A., Nagaraja, H., & Miser, W. F. (2012). Comparative effectiveness of a mindful eating intervention to a diabetes self-management intervention among adults with type 2 diabetes: a pilot study. *Journal of the Academy of Nutrition and Dietetics*, 112(11), 1835-1842.
- Miller, C. K., Kristeller, J. L., Headings, A., & Nagaraja, H. (2014). Comparison of a mindful eating intervention to a diabetes self-management intervention among adults with type 2 diabetes: a randomized controlled trial. *Health Education & Behavior*, 41(2), 145-154.

## SELF-REPORT MINDFULNESS VALIDITY

- Moore, R. C., Depp, C. A., Wetherell, J. L., & Lenze, E. J. (2016). Ecological momentary assessment versus standard assessment instruments for measuring mindfulness, depressed mood, and anxiety among older adults. *Journal of Psychiatric Research*, 75, 116-123.
- Morone, N. E., Rollman, B. L., Moore, C. G., Li, Q., & Weiner, D. K. (2009). A mind-body program for older adults with chronic low back pain: Results of a pilot study. *Pain Medicine*, 10(8), 1395-1407.
- Morone, N. E., Greco, C. M., Moore, C. G., Rollman, B. L., Lane, B., Morrow, L. A., ... & Weiner, D. K. (2016). A mind-body program for older adults with chronic low back pain: A randomized clinical trial. *JAMA Internal Medicine*, 176(3), 329-337.
- Nakamura, Y., Lipschitz, D. L., Kuhn, R., Kinney, A. Y., & Donaldson, G. W. (2013). Investigating efficacy of two brief mind-body intervention programs for managing sleep disturbance in cancer survivors: A pilot randomized controlled trial. *Journal of Cancer Survivorship*, 7(2), 165-182.
- Perich, T., Manicavasagar, V., Mitchell, P. B., Ball, J. R., & Hadzi-Pavlovic, D. (2013). A randomized controlled trial of mindfulness-based cognitive therapy for bipolar disorder. *Acta Psychiatrica Scandinavica*, 127(5), 333-343.
- Polusny, M.A., Erbes, C.R., Thuras, P., Moran, A., Lamberty, G.J., Collins, R.C.,...& Lim, K.O. (2015). Mindfulness-based stress reduction for posttraumatic stress disorder among veterans: A randomized clinical trial. *JAMA*, 314(5), 456-465.
- Possemato, K., Bergen - Cico, D., Treatman, S., Allen, C., Wade, M., & Pigeon, W. (2016). A randomized clinical trial of primary care brief mindfulness training for veterans with PTSD. *Journal of Clinical Psychology*, 72(3), 179-193.
- Pradhan, E. K., Baumgarten, M., Langenberg, P., Handwerker, B., Gilpin, A. K., Magyari, T., ... & Berman, B. M. (2007). Effect of mindfulness-based stress reduction in rheumatoid arthritis patients. *Arthritis Care & Research*, 57(7), 1134-1142.
- Schmidt, S., Grossman, P., Schwarzer, B., Jena, S., Naumann, J., & Walach, H. (2011). Treating fibromyalgia with mindfulness-based stress reduction: results from a 3-armed randomized controlled trial. *Pain*, 152(2), 361-369.
- Segal, Z. V., Bieling, P., Young, T., MacQueen, G., Cooke, R., Martin, L., ... & Levitan, R. D. (2010). Antidepressant monotherapy vs sequential pharmacotherapy and mindfulness-based cognitive therapy, or placebo, for relapse prophylaxis in recurrent depression. *Archives of General Psychiatry*, 67(12), 1256-1264.
- Shahar, B., Britton, W. B., Sbarra, D. A., Figueredo, A. J., & Bootzin, R. R. (2010). Mechanisms of change in mindfulness-based cognitive therapy for depression: Preliminary evidence from a randomized controlled trial. *International Journal of Cognitive Therapy*, 3(4), 402-418.

## SELF-REPORT MINDFULNESS VALIDITY

Strauss, C., Hayward, M., & Chadwick, P. (2012). Group person-based cognitive therapy for chronic depression: A pilot randomized controlled trial. *British Journal of Clinical Psychology*, 51(3), 345-350.

Van Aalderen, J. R., Donders, A. R. T., Giommi, F., Spinhoven, P., Barendregt, H. P., & Speckens, A. E. M. (2012). The efficacy of mindfulness-based cognitive therapy in recurrent depressed patients with and without a current depressive episode: A randomized controlled trial. *Psychological Medicine*, 42(05), 989-1001.

Vøllestad, J., Sivertsen, B., & Nielsen, G. H. (2011). Mindfulness-based stress reduction for patients with anxiety disorders: Evaluation in a randomized controlled trial. *Behaviour Research and Therapy*, 49(4), 281-288.

Witkiewitz, K., Greenfield, B. L., & Bowen, S. (2013). Mindfulness-based relapse prevention with racial and ethnic minority women. *Addictive Behaviors*, 38(12), 2821-2824.

Wong, S. Y. S., Yip, B. H. K., Mak, W. W. S., Mercer, S., Cheung, E. Y. L., Ling, C. Y. M., ... & Lee, T. M. C. (2016). Mindfulness-based cognitive therapy v. group psychoeducation for people with generalised anxiety disorder: Randomised controlled trial. *The British Journal of Psychiatry*, 209(1), 68-75.

Zgierska, A. E., Burzinski, C. A., Cox, J., Kloke, J., Stegner, A., Cook, D. B., ... & Bačkonja, M. (2016). Mindfulness meditation and cognitive behavioral therapy intervention reduces pain severity and sensitivity in opioid-treated chronic low back pain: Pilot findings from a randomized controlled trial. *Pain Medicine*, 17(10), 1865-1881.

## SELF-REPORT MINDFULNESS VALIDITY

Supplemental Materials Table 6. Clinical outcome measures included

Study	Outcome
Alberts 2012	Body Mass Index
Alberts 2012	Dutch Eating Behaviour Questionnaire
Alberts 2012	Body Shape Questionnaire
Alberts 2012	General Food Craving Questionnaire Trait
Alexander 2012	Beck Depression Inventory
Atkinson 2016	Eating Disorder Examination Questionnaire
Atkinson 2016	Dutch Eating Behaviour Questionnaire
Atkinson 2016	Sociocultural Attitudes Towards Appearance Scale
Atkinson 2016	Eating Disorder Examination Questionnaire
Atkinson 2016	Clinical Impairment Assessment - Psychosocial impairment related to eating
Bedard 2014	Beck Depression Inventory II
Bedard 2014	Patient Health Questionnaire - 9
Bedard 2014	Symptom Checklist-90R - Depression
Bowen 2009	Alcohol and drug use days
Bowen 2009	Short Inventory of Problems
Bowen 2009	Penn Alcohol Craving Scale
Bowen 2009	Addiction Severity Index
Brewer 2009	Cocaine use days
Brewer 2009	Alcohol use days
Brewer 2009	Drug craving
Brown 2013	Pain Stages of Change Questionnaire
Brown 2013	Pain Attitudes Questionnaire - Perceived control
Brown 2013	Short-Form McGill Pain Questionnaire - Sensory Pain
Brown 2013	Short-Form McGill Pain Questionnaire - Affective Pain
Brown 2013	Laser pain unpleasantness
Cathcart 2014	Muscle Tenderness
Cathcart 2014	Pain Detection Threshold at Finger
Cathcart 2014	Pain Detection Threshold at Shoulder
Cathcart 2014	Cold-pain detection threshold
Cathcart 2014	Pain rating at 10 sec of immersion
Cathcart 2014	Pain rating at 20 sec of immersion
Cathcart 2014	Cold pain tolerance in seconds
Cathcart 2014	Conditioned Pain Modulation
Cathcart 2014	Headache frequency
Cathcart 2014	Headache duration
Cathcart 2014	Headache intensity
Chadwick 2009	Southampton Mindfulness Voices Questionnaire
Chadwick 2009	Beliefs about Voices Questionnaire Revised
Chadwick 2009	Psychiatric symptom rating scale - auditory hallucinations



## SELF-REPORT MINDFULNESS VALIDITY

---

Chadwick 2009	Psychiatric symptom rating scale - paranoia
Chavooshi 2016	Numerical Pain Rating Scale
Cherkin 2016	Roland Disability Questionnaire (pain)
Cherkin 2016	Pain bothersome rating
Cherkin 2016	Characteristic pain intensity
Cherkin 2016	Global Improvement in Pain
Cherkin 2016	Pain Catastrophizing Scale
Cherkin 2016	Pain Self-Efficacy Questionnaire
Cherkin 2016	Chronic Pain Acceptance Questionnaire - Total score
Chiesa 2015	Hamilton Rating Scale for Depression
Chiesa 2015	Beck Depression Inventory
Colgan 2016	PTSD Checklist
Daubenmier 2011	Restrained eating
Daubenmier 2011	Emotional eating
Daubenmier 2011	External eating
Daubenmier 2011	Weight
Daubenmier 2011	Abdominal fat
Daubenmier 2011	Trunk/leg fat ratio
Davis 2013	2-week point prevalence abstinence
Davis 2013	Days abstinent two-weeks post-quit
Davis 2014	4-week point prevalence abstinence
Davis 2014	24-week point prevalence abstinence
Davis 2014	Urge ratings
Day 2014	Headache frequency
Day 2014	Headache duration
Day 2014	Headache index
Day 2014	Peak headache intensity
Day 2014	Headache distress
Day 2014	Headache disability, equivalent hours
Day 2014	Acetaminophen
Day 2014	Morphine
Day 2014	Brief Pain Inventory - intensity
Day 2014	Brief Pain Inventory - interference
Day 2014	Pain Catastrophizing Scale
Day 2014	Chronic Pain Acceptance Questionnaire
Day 2014	Headache Management Self-Efficacy Scale
Eisendrath 2016	Multivariate depression score % reduction
Eisendrath 2016	Treatment responder % based on Hamilton Rating Scale for Depression
Eisendrath 2016	Remission rates %
Eisendrath 2016	Ruminative Response Scale
Garland 2010	Alcohol Craving
Garland 2010	Impaired Alcohol Response Inhibition
Garland 2016	Penn Alcohol Craving Scale

---

## SELF-REPORT MINDFULNESS VALIDITY

---

Goldin 2016	Liebowitz Social Anxiety Scale - Self-report
Goldin 2016	Subtle Avoidance Frequency Examination
Hanstede 2008	Obsessive Compulsive Disorder complaints
Hepark 2015	Conners' Adult ADHD Rating Scale - Investigator Rating
Hepark 2015	Conners' Adult ADHD Rating Scale - Self Report
Hepark 2015	Behavior Rating Inventory of Executive Function - Adult
Hepark 2015	False alarms to NoGo stimuli
Hepark 2015	Correct hits to Go stimuli
Hepark 2015	Correctly rejected NoGo stimuli
Hepark 2015	False alarm reaction time
Hepark 2015	Correct hits to Go stimuli reaction time
Hoge 2013	Hamilton Anxiety Rating Scale
Hoge 2013	Clinical Global Impressions - Severity of Illness
Hoge 2013	Beck Anxiety Inventory
Hoge 2013	Pittsburgh Sleep Quality Index
Hoge 2013	Trier Social Stress Test – State-Trait Anxiety Inventory
Hoge 2013	Trier Social Stress Test – Subjective Units of Distress
Hoge 2013	Trier Social Stress Test - Positive Self-Statements During Public Speaking
Hoge 2013	Trier Social Stress Test - Negative Self-Statements During Public Speaking
Hoge 2013	Penn State Worry Questionnaire
Imani 2015	Addiction Severity Index - Alcohol subscale
Imani 2015	Addiction Severity Index - Opium subscale
Imani 2015	Positive morphine urine tests
Jazaieri 2012	Sheehan Disability Scale
Jazaieri 2012	Negative Self-Endorsement
Jazaieri 2012	Positive Self-Endorsement
Jazaieri 2012	Negative emotion reactivity
Jazaieri 2012	Negative emotion regulation
Jazaieri 2012	Liebowitz Social Anxiety Scale
Jazaieri 2012	Social Interaction Anxiety Scale
Jazaieri 2012	Social anxiety clinical symptoms
Kearney 2013	PTSD Checklist
Keune 2011	Beck Depression Inventory
Keune 2011	Response Styles Questionnaire - Symptom-focused rumination
Keune 2011	Response Styles Questionnaire - Self-focused rumination
Kocovski 2013	Social Phobia Inventory
Kocovski 2013	Liebowitz Social Anxiety Scale - Clinician report
Kocovski 2013	Clinical Global Impression Severity - Clinician report
Koszycki 2016	Liebowitz Social Anxiety Scale
Koszycki 2016	Social Phobia Inventory
Koszycki 2016	Clinical Global Impression - Severity Scale
Kuyken 2008	Relapse of depression
Kuyken 2008	Hamilton Rating Scale for Depression

---

## SELF-REPORT MINDFULNESS VALIDITY

---

Kuyken 2008	Beck Depression Inventory - II
Kuyken 2008	Sad mood pre- to post-induction
Kuyken 2008	Sad mood post-induction
Kuyken 2008	Dysfunctional Attitudes Scale pre- to post-induction
Kuyken 2008	Dysfunctional Attitudes Scale post-induction
Kuyken 2008	Sad mood pre-induction
Kuyken 2008	Dysfunctional Attitudes Scale pre-induction
Langer 2012	Clinical Global Impression - Schizophrenia
Lopez-Navarro2015	Positive and Negative Syndrome Scale Positive Symptoms
Lopez-Navarro 2015	Positive and Negative Syndrome Scale Negative Symptoms
Lopez-Navarro 2015	Positive and Negative Syndrome Scale General Symptoms
Manicavasgar 2011	Beck Depression Inventory - II
Manicavasgar 2011	Rumination Response Style Questionnaire
Mann 2016	Beck Depression Inventory II
McIndoo 2016	Beck Depression Inventory - II
McIndoo 2016	Hamilton Rating Scale for Depression
McIndoo 2016	Ruminative Response Scale
Miller 2014	Waist circumference
Miller 2014	Body Mass Index
Miller 2014	Metabolic equivalent hour/week
Miller 2014	Weight
Miller 2014	Energy (calories)
Miller 2014	Eating self-efficacy
Miller 2014	Three-factor eating questionnaire - control/restraint
Miller 2014	Three-factor eating questionnaire - disinhibition of control
Miller 2014	Three-factor eating questionnaire - hunger susceptibility
Moore 2016	NIH PROMIS Electronic Momentary Assessment Anxiety
Moore 2016	NIH PROMIS paper-and-pencil anxiety
Morone 2009	Roland Disability Questionnaire
Morone 2009	Short Form-36 Pain Score
Morone 2009	Chronic Pain Self-Efficacy Scale: Pain Self-Efficacy
Morone 2009	McGill Pain Questionnaire: Total Score
Morone 2016	Roland Disability Questionnaire
Morone 2016	Numeric Pain Rating Scale - Average
Morone 2016	Numeric Pain Rating Scale - Current
Morone 2016	Numeric Pain Rating Scale - Most Severe
Morone 2016	Chronic Pain Self-Efficacy Scale - Coping
Morone 2016	Chronic Pain Self-Efficacy Scale - Function
Morone 2016	Chronic Pain Self-Efficacy Scale - Pain
Morone 2016	Pain Catastrophizing Scale of Coping Questionnaire
Nakamura 2013	Sleep Problems Index - II
Nakamura 2013	Medical Outcome Sleep Study Scale - Sleep problems index II
Perich 2013	Montgomery-Asberg Depression Scale

---

## SELF-REPORT MINDFULNESS VALIDITY

---

Perich 2013	Young Mania Rating Scale
Perich 2013	Depression Anxiety Stress Scales - Depression
Perich 2013	Response Style Questionnaire - Rumination
Perich 2013	Response Style Questionnaire - Dangerous
Perich 2013	Mania relapse
Perich 2013	Depressive relapse
Polusny 2015	PTSD Checklist
Polusny 2015	Clinician- Administered PTSD Scale
Polusny 2015	Loss of PTSD diagnosis
Possemato 2016	Clinician-Administered PTSD Scale
Possemato 2016	PTSD Checklist
Pradhan 2007	Disease Activity Score in 28 Joints
Schmidt 2011	Activity - 9-12
Schmidt 2011	Activity - 12-15
Schmidt 2011	Activity - 15-18
Schmidt 2011	Activity - 19-21
Schmidt 2011	Health Related Quality of Life
Schmidt 2011	Fibromyalgia Impact Questionnaire
Schmidt 2011	Pain Perception Scale - Affective
Schmidt 2011	Pain Perception Scale - Sensory
Segal 2010	Experiences Questionnaire - Rumination
Segal 2010	Relapse rates
Shahar 2010	Beck Depression Inventory
Shahar 2010	Ruminative Response Scale - Brooding
Shahar 2010	Ruminative Response Scale - Reflection
Strauss 2012	Beck Depression Inventory -II
VanAalderen 2012	Hamilton Rating Scale of Depression
VanAalderen 2012	Beck Depression Inventory
VanAalderen 2012	Rumination on Sadness Scale
Vollestad 2011	Beck Anxiety Inventory
Vollestad 2011	Penn State Worry Questionnaire
Vollestad 2011	State-Trait Anxiety Inventory-State
Vollestad 2011	State-Trait Anxiety Inventory -Trait
Vollestad 2011	Bergen Insomnia Scale
Wells 2014	Headache Impact Test
Wells 2014	Migraine Disability Assessment
Wells 2014	Headache Management Self-Efficacy Scale
Wells 2014	Migraine-Specific Quality of Life
Wong 2016	Beck Anxiety Inventory
Wong 2016	Penn State Worry Questionnaire
Zgierska 2016	Brief Pain Inventory
Zgierska 2016	Oswestry Disability Index total score
Zgierska 2016	Morphine Equivalent Dose mg per day

---

## SELF-REPORT MINDFULNESS VALIDITY

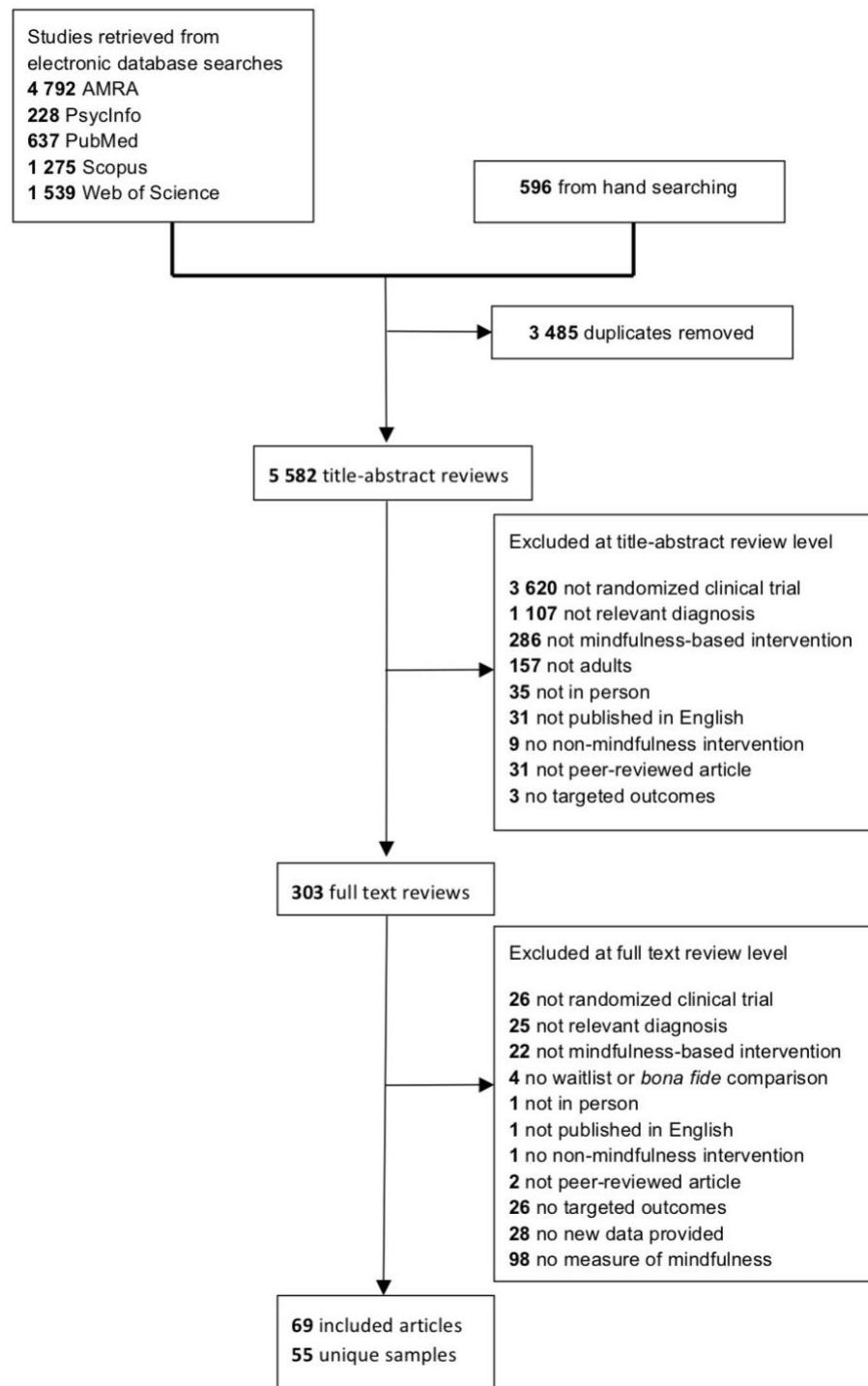
---

Zgierska 2016	Chronic Pain Acceptance Questionnaire
Zgierska 2016	Pain Intensity
Zgierska 2016	Pain Unpleasantness

---

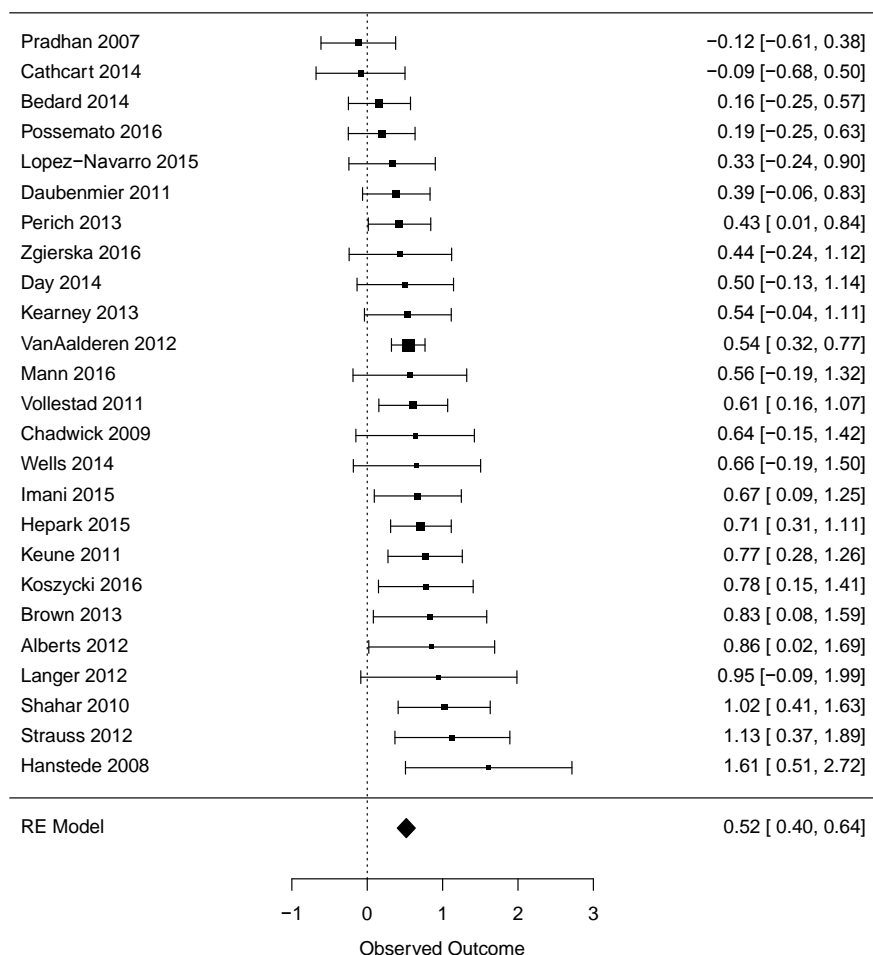
Note: NIH PROMIS = National Institute of Health Patient-Reported Outcomes Measurement Information System.

## SELF-REPORT MINDFULNESS VALIDITY



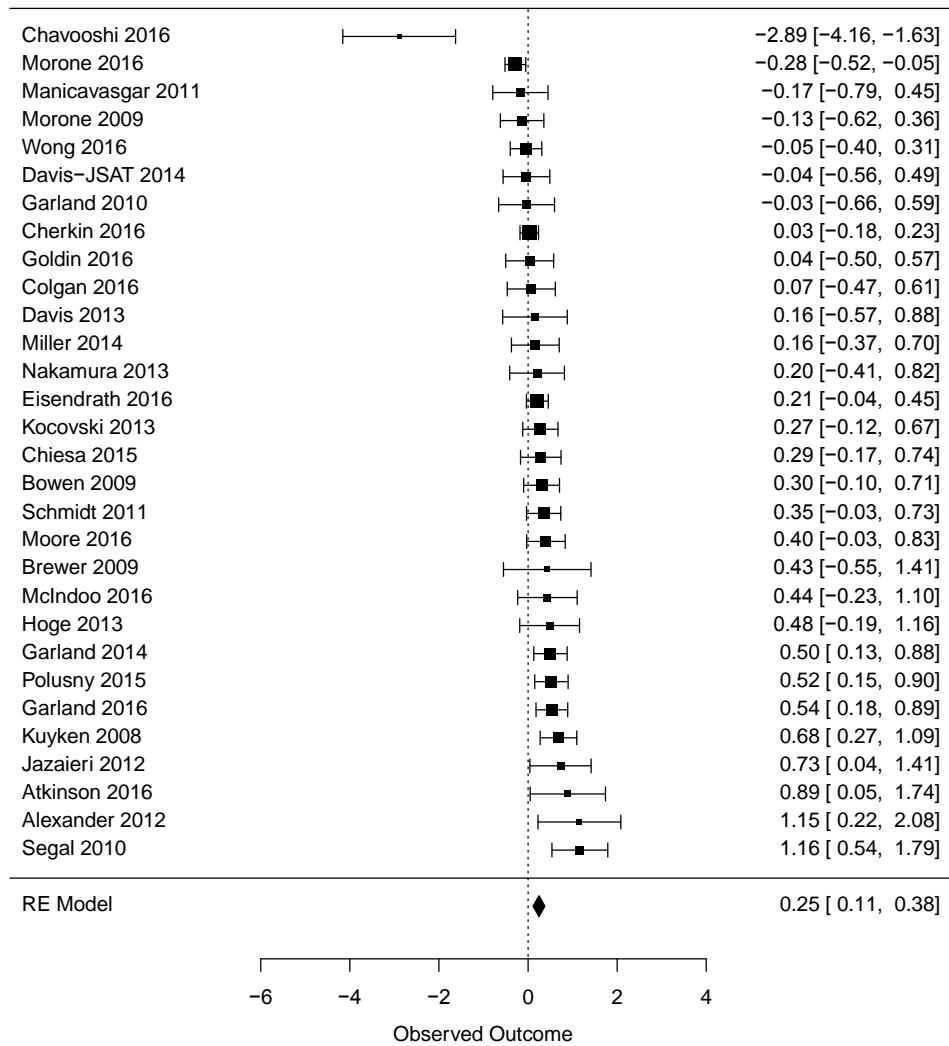
Supplemental Materials Figure 1. Summary of literature search

## SELF-REPORT MINDFULNESS VALIDITY



Supplemental Materials Figure 2. Forest plot of pre-post relative responsiveness (mindfulness versus waitlist conditions) on measures of mindfulness.

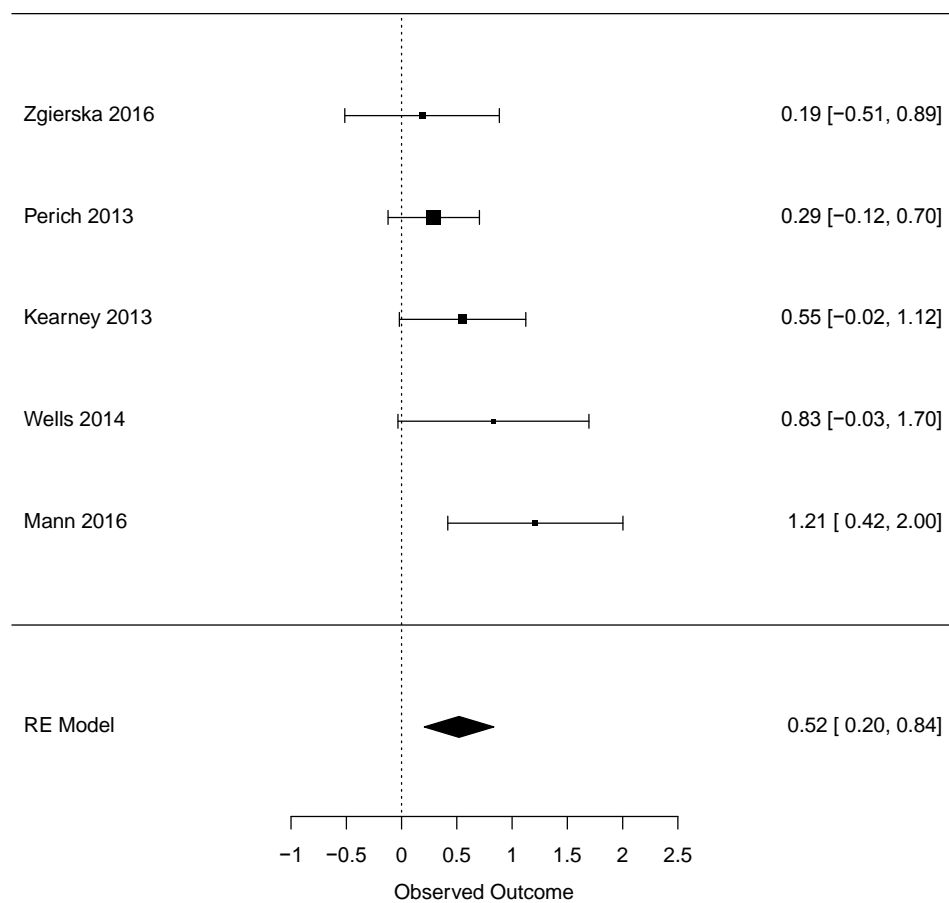
## SELF-REPORT MINDFULNESS VALIDITY



Supplemental Materials Figure 3. Forest plot of pre-post relative responsiveness (mindfulness versus *bona fide* comparison condition) on measures of mindfulness.

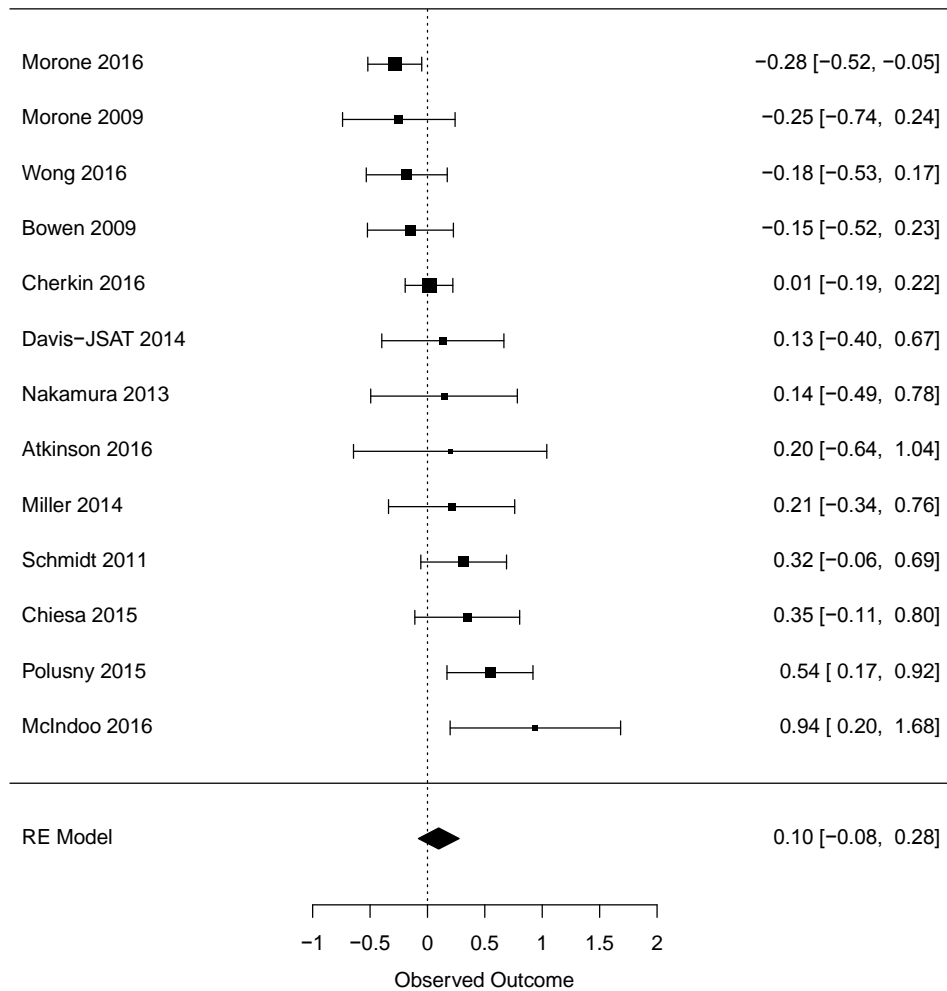


## SELF-REPORT MINDFULNESS VALIDITY

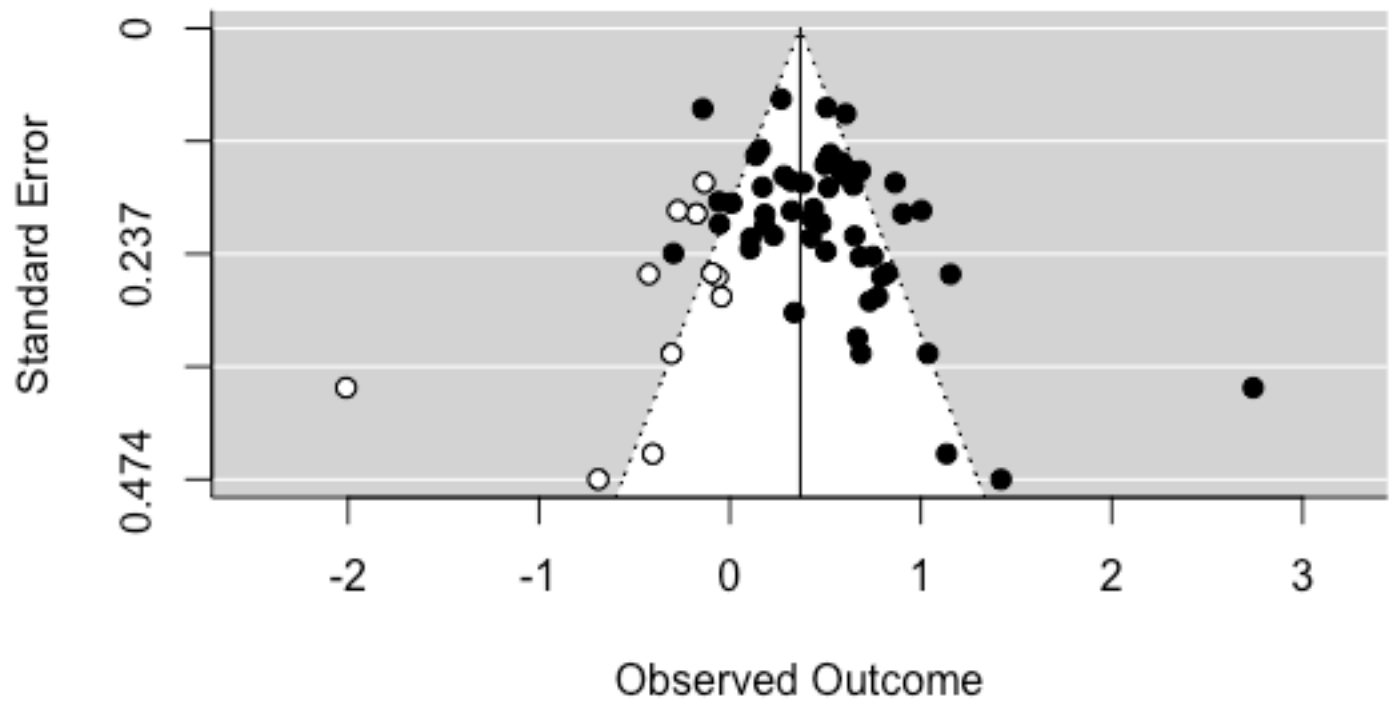


Supplemental Materials Figure 4. Forest plot of pre- to follow-up relative responsiveness (mindfulness versus waitlist conditions) on measures of mindfulness.

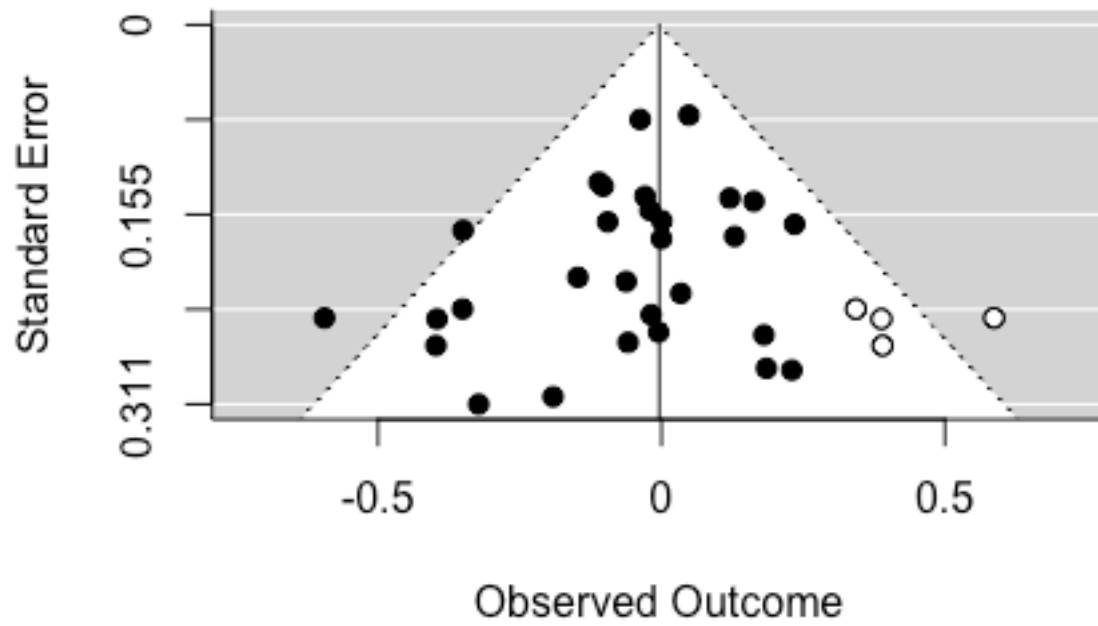
## SELF-REPORT MINDFULNESS VALIDITY



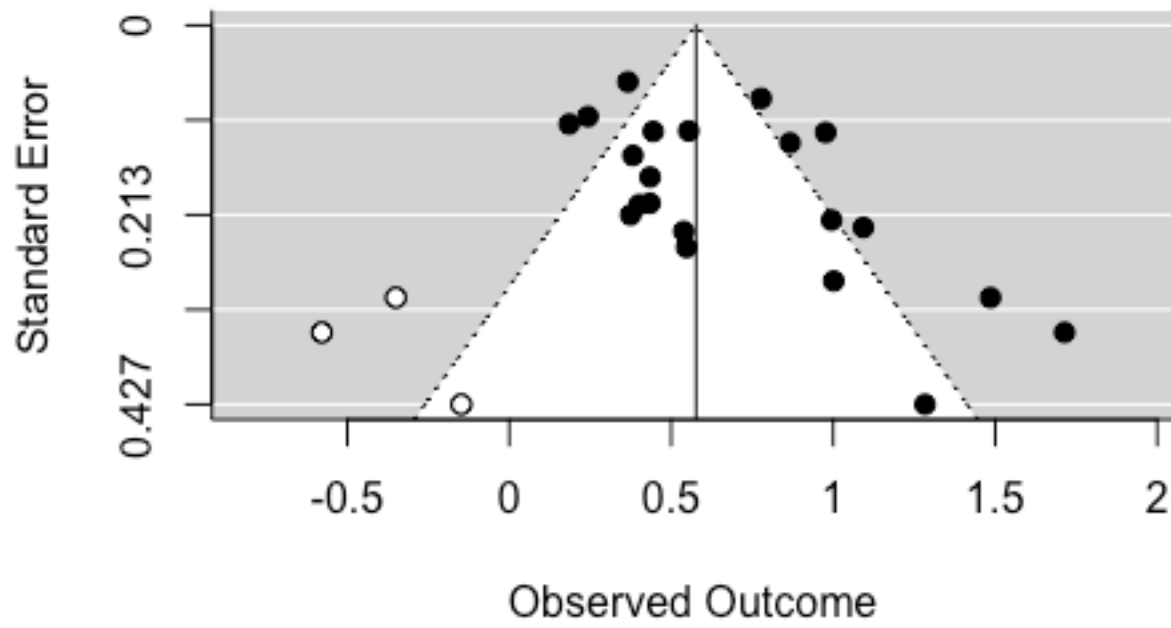
Supplemental Materials Figure 5. Forest plot of pre- to follow-up relative responsiveness (mindfulness versus *bona fide* comparison condition) on measures of mindfulness.



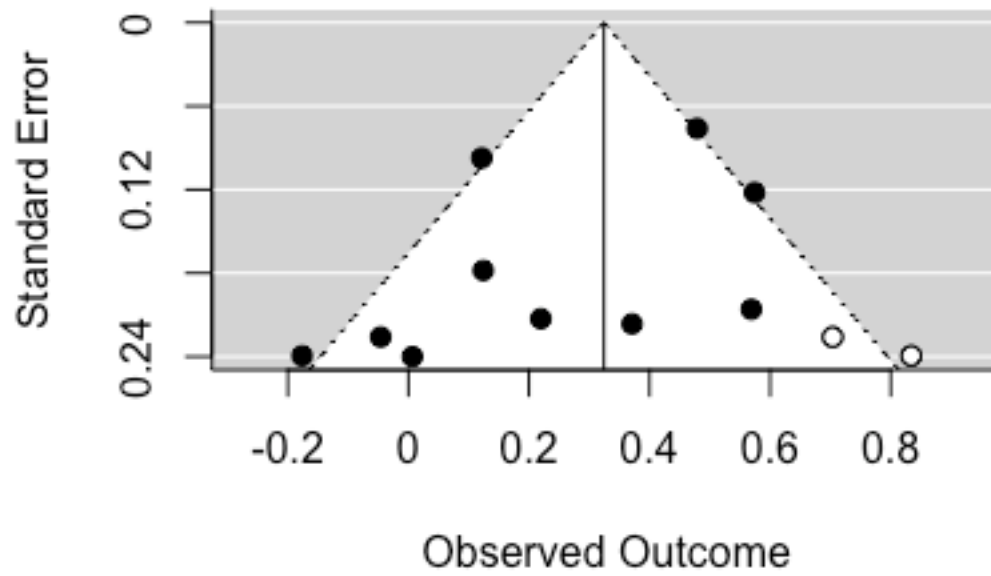
Supplemental Materials Figure 6. Funnel plot with imputed studies for pre-post within-group changes on mindfulness in mindfulness conditions.



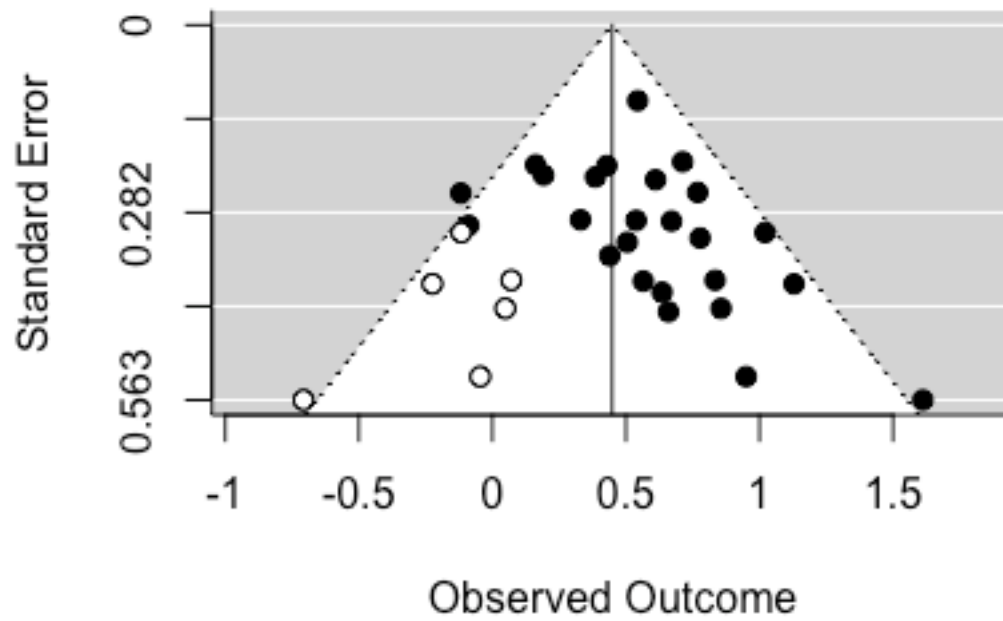
Supplemental Materials Figure 7. Funnel plot with imputed studies for pre-post within-group changes on mindfulness in waitlist control conditions.



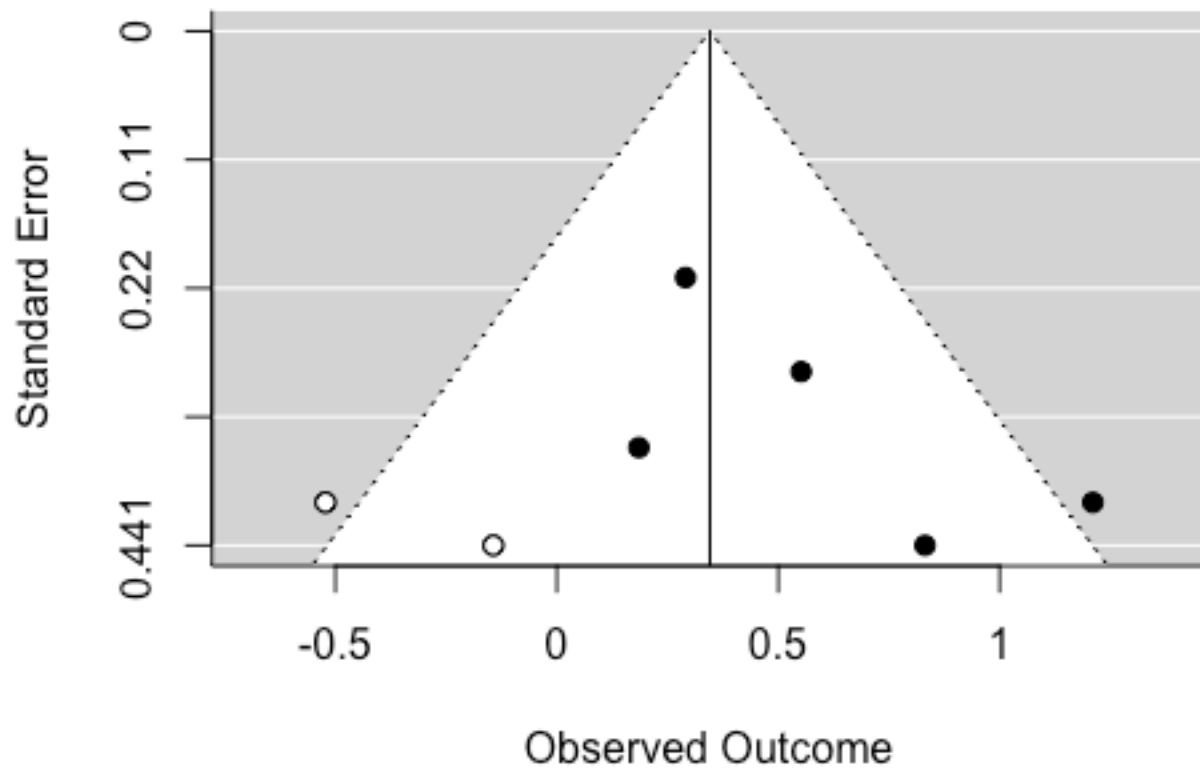
Supplemental Materials Figure 8. Funnel plot with imputed studies for pre- to follow-up within-group changes on clinical outcomes in mindfulness conditions.



Supplemental Materials Figure 9. Funnel plot with imputed studies for pre- to follow-up within-group changes on clinical outcomes in waitlist control conditions.

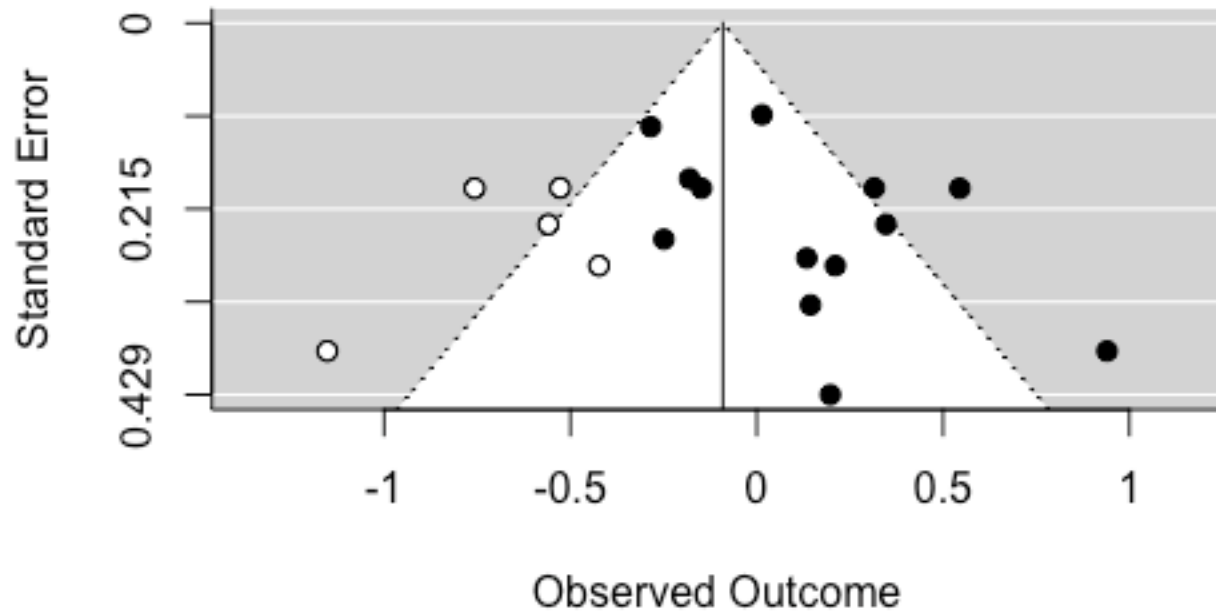


Supplemental Materials Figure 10. Funnel plot with imputed studies for pre-post relative responsiveness on mindfulness in mindfulness versus waitlist control conditions.

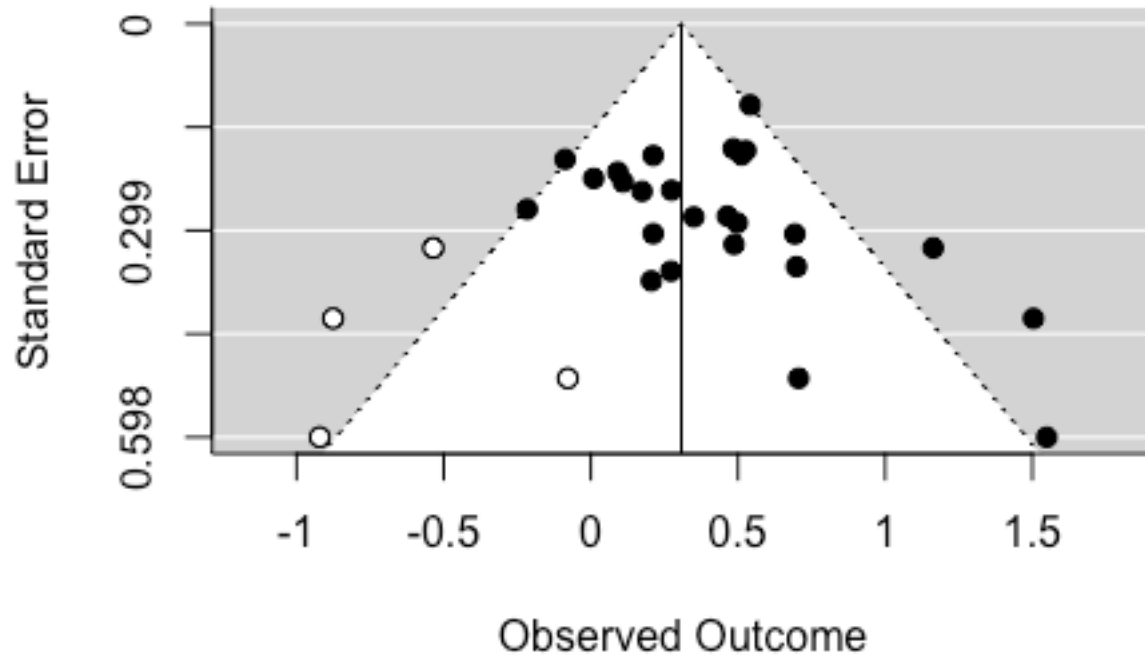


Supplemental Materials Figure 11. Funnel plot with imputed studies for pre- to follow-up relative responsiveness on mindfulness in mindfulness versus waitlist control conditions.

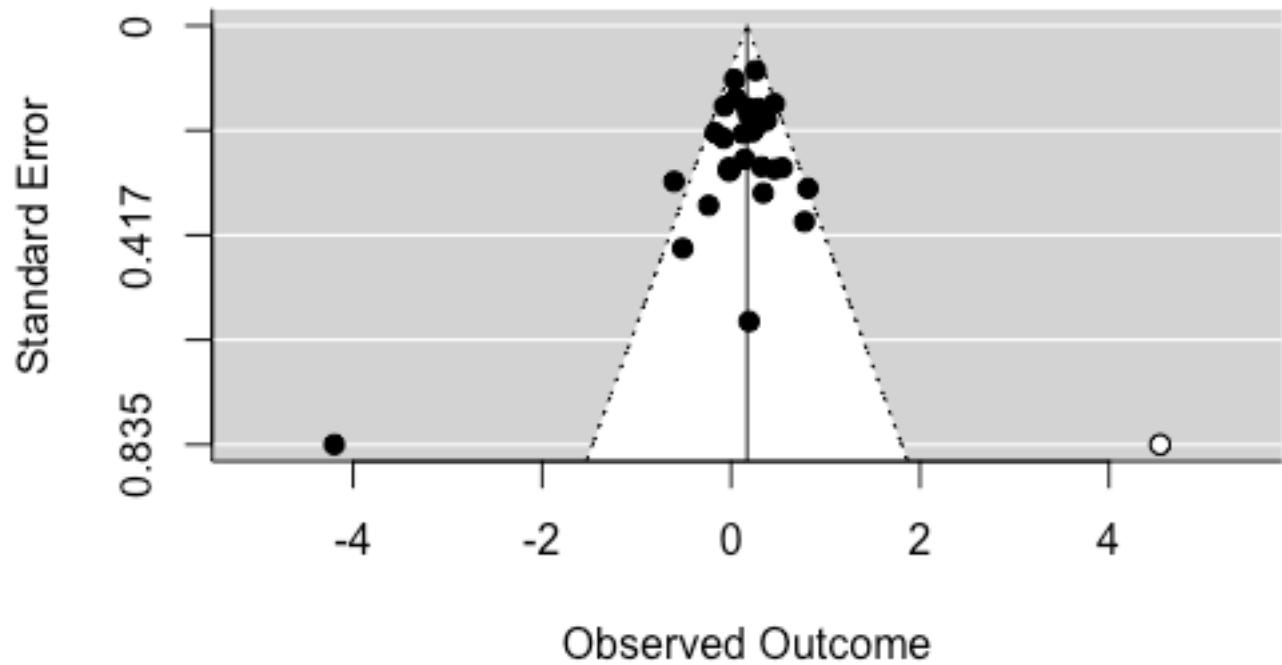




Supplemental Materials Figure 12. Funnel plot with imputed studies for pre- to follow-up relative responsiveness on mindfulness in mindfulness versus *bona fide* comparison conditions.



Supplemental Materials Figure 13. Funnel plot with imputed studies for pre-post relative responsiveness on clinical outcomes in mindfulness versus waitlist control conditions.



Supplemental Materials Figure 14. Funnel plot with imputed studies for pre-post relative responsiveness on clinical outcomes in mindfulness versus *bona fide* comparison conditions.