Psychotherapy Research, 2020
Vol. 30, No. 6, 753–762, https://doi.org/10.1080/10503307.2019.1664783

EMPIRICAL PAPER

Allegiance effects in mindfulness-based interventions for psychiatric disorders: A meta-re-analysis

SIMON B. GOLDBERG [1,2] & RAYMOND P. TUCKER [3,4]

[1]*Department of Counseling Psychology, University of Wisconsin–Madison, Madison, WI, USA;* [2]*Center for Healthy Minds, University of Wisconsin–Madison, Madison, WI, USA;* [3]*Department of Psychology, Louisiana State University, Baton Rouge, LA, USA &* [4]*Department of Psychiatry, Louisiana State University Health Sciences Center, Baton Rouge, LA, USA*

(*Received 29 June 2019; revised 27 August 2019; accepted 2 September 2019*)

Abstract

Objectives: A recent meta-analysis reported that mindfulness-based interventions (MBIs) outperform specific active control conditions but not evidence-based treatments (EBTs) across various psychiatric conditions. Given both comparison conditions represent *bona fide* treatments, the superiority of MBIs over other *bona fide* treatments is unexpected. The current study examined researcher allegiance (RA) as a potential source of bias that may explain this result. **Method:** All studies from the original meta-analysis that compared MBIs with *bona fide* psychological treatments were included. RA was independently coded using established methods. A series of meta-analyses examined the RA-outcome association and the degree to which RA may account for the effect of EBT status. **Results:** Sixty independent comparisons ($n = 5{,}627$) were included. MBIs outperformed *bona fide* treatment comparisons overall ($g = 0.13$), but effects were smaller with EBT comparisons. RA towards MBIs was associated with larger effects. No evidence for superiority of MBIs was found when RA was absent or balanced. Further, EBT status no longer predicted effects when controlling for RA. **Conclusions:** RA appears to be a potential source of bias in MBI research that should be considered when interpreting existing studies (clinical trials, meta-analyses) and planning future studies. RA may account for smaller effects when using EBT comparisons.

Keywords: researcher allegiance; mindfulness-based interventions; psychiatric disorders; meta-analysis; relative efficacy

Clinical and Methodological Significance: The current study is the first to meta-analytically examine researcher allegiance as a source of bias within mindfulness research. Results support the notion that researchers' belief in the superiority of mindfulness and methodological choices that may follow from this belief (e.g., providing greater supervision or training for the mindfulness condition relative to the control condition) may produce larger effects in favour of mindfulness. This highlights the importance of comparing mindfulness-based interventions with frontline, evidence-based therapies and including adversarial collaborators when seeking to test relative efficacy.

Public and scientific interest in mindfulness meditation has grown dramatically in the past several decades (Black, 2014; Wieczner, 2016). Between 2012 and 2017, use of meditation in the United States tripled, with 14.2% of adults reporting past year use (Clarke, Barnes, Black, Stussman, & Nahin, 2018). Since the appearance of Mindfulness-Based Stress Reduction (MBSR; Kabat-Zinn, 1990), numerous mindfulness-based interventions (MBIs) have been developed and targeted towards reducing psychiatric symptoms. For example, Mindfulness-Based Cognitive Therapy (MBCT; Segal, Williams, & Teasdale, 2013) was developed to prevent relapse in major depressive disorder. MBCT was recently been granted evidence-based treatment (EBT) status by the American Psychological Association's (APA) Society of Clinical Psychology (Division 12; Society of Clinical Psychology, 2019) as a treatment for depression. Other MBIs have been developed for treating substance use (e.g., Mindfulness-Based Relapse Prevention, Mindfulness Oriented Recovery Enhancement; Bowen et al., 2014; Garland et al., 2014) and eating disorders (e.g., Mindfulness-Based Eating Awareness

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

Training; Kristeller & Wolever, 2010), among other psychiatric conditions.

There are over a hundred randomized clinical trials (RCTs) testing MBIs for psychiatric disorders. In a recent meta-analysis, Goldberg et al. (2018) aggregated the effects of MBIs on psychiatric symptoms in clinical populations (e.g., depression, anxiety, substance use). Control conditions were coded using a five-tier system, with two conditions – specific active controls, EBTs – representing treatments intended to be therapeutic (i.e., *bona fide* treatments per Wampold et al.'s [1997] definition). Of note, not all of the treatments that Goldberg et al. coded as *bona fide* were traditional forms of psychotherapy (e.g., health education programme for anxiety; Moore, Depp, Wetherell, & Lenze, 2016). However, according to Goldberg et al., all treatments coded as *bona fide* did include components that were intended to be therapeutic (e.g., Moore et al. [2016]'s health education programme was based on Lorig et al.'s [2012] self-management programme for chronic conditions including depression and anxiety). This broader notion of *bona fide* treatments (i.e., beyond traditional forms of psychotherapy) is arguably appropriate given the contexts in which MBIs are sometimes delivered and tested (e.g., non-psychiatric medical settings in which health education and similar programmes may reflect the standard of care; Bohlmeijer, Prenger, Taal, & Cuijpers, 2010). Further, consistent with Wampold et al. (1997), Goldberg et al. coded treatments as *bona fide* if they included components drawn from *bona fide* treatments (i.e., treatments were considered *bona fide* even if they were modified to omit content that might overlap with the MBI, e.g., Moore et al. [2016] removing relaxation and meditation strategies from Lorig et al. [2012]). In contrast, Goldberg et al. coded treatments as non-specific active controls if they were deemed not intended to be therapeutic. Examples of this include watching short films on social and cultural topics (Langer, Cangas, & Gallego, 2010) or reminiscing about personal experiences (Helmes & Ward, 2017).

Across 142 non-overlapping samples ($n = 12{,}005$), Goldberg et al. found that MBIs outperformed no treatment ($d = 0.55$), minimal treatment ($d = 0.37$), non-specific active controls ($d = 0.35$), and specific active controls ($d = 0.23$), and performed on par ($d = -0.004$) with treatments at the time defined as EBTs by APA's Division 12 and similar organizations (e.g., American Lung Association for smoking cessation treatments). The authors conclude that MBIs may be a promising frontline approach for a variety of psychiatric conditions.

While perhaps accurately highlighting the therapeutic potential of MBIs, Goldberg et al.'s (2018) results include an unexpected finding: superiority of MBIs over specific active controls that were intended to be therapeutic (e.g., progressive muscle relaxation for chronic pain; Schmidt et al., 2011). Although small in magnitude ($d = 0.23$), the larger effects observed in MBIs relative to other *bona fide* treatments flies in the face of a substantial body of evidence that suggests *bona fide* treatments work equally well (e.g., Baardseth et al., 2013; Imel, Wampold, Miller, & Fleming, 2008; Kivlighan et al., 2015; Wampold et al., 1997; Wampold & Imel, 2015). When the comparison groups were EBTs (e.g., cognitive behavioural therapy for anxiety; Arch et al., 2013), the comparison with MBIs ($d = -0.004$) showed the more expected pattern of equivalence (i.e., uniform efficacy; Wampold & Imel, 2015). Goldberg et al. (2018) evaluated several sources of bias that could contribute to this unexpected finding: failure to report of intent-to-treat, exclusive reliance on self-report measures, lack of treatment dosage matching between conditions. However, none of these factors tended to moderate between-group effects.

One source of bias that could explain this pattern of findings but was not explicitly addressed by Goldberg et al. (2018) is researcher allegiance (RA).[1] RA has been defined as a researcher's "belief in the superiority of a treatment [and] … the superior validity of the theory of change that is associated with the treatment" (Leykin & DeRubeis, 2009, p. 55). The potentially potent (and perhaps pernicious) impact of RA has been recognized within psychotherapy research for decades (e.g., Berman, Miller, & Massman, 1985). RA may manifest in more or less obvious ways, but can include a researcher serving as the therapist or supervisor in a study testing a therapy they developed or advocate for, the therapists in one condition receiving more training or supervision, and a greater dosage of treatment being provided to patients in one condition (Munder, Gerger, Trelle, & Barth, 2011; Yulish et al., 2017). The RA-outcome association is robust: Munder, Brütsch, Leonhart, Gerger, and Barth (2013) conducted a meta-meta-analysis (i.e., review of reviews), demonstrating a small-to-moderate magnitude RA-outcome association ($r = .26$) across 30 meta-analyses. That is to say, results of clinical trials tended to favour the condition to which researchers had a stronger allegiance (i.e., believed in the superiority of the treatment, methodologically favoured the treatment by providing more supervision or training, etc.). As further testament to the power of RA, the RA-outcome association was larger in meta-analyses conducted by researchers allegiant to the RA hypothesis. Some have argued that the causal direction linking RA and outcome flows from outcome to RA, with

researchers showing allegiance to better performing therapies (i.e., an "epiphenomenon of true efficacy differences," Munder, Flückiger, Gerger, Wampold, & Barth, 2012, p. 631). However, the appearance of the RA-outcome association when restricting to treatments of equivalent efficacy suggests this is not the case (Munder et al., 2012). Arguably, the influence of RA outside of the context of psychotherapy research has been evident in recent years in the "replication crisis" (Maxwell, Lau, & Howard, 2015) and the associated movement towards increased transparency in scientific research (Open Science Collaboration, 2012, 2015). There has been increasing recognition (and criticism) of questionable research practices (e.g., selective reporting of outcomes, "p-hacking"; Head, Holman, Lanfear, Kahn, & Jennions, 2015) as mechanisms through which RA may influence outcomes.

As applied to the case of Goldberg et al. (2018), it is possible that the apparent superiority of MBIs to specific active control and equivalence to EBTs is actually due to RA, not due to true superiority of MBIs over other *bona fide* therapies nor to the fact that EBTs are necessarily a more efficacious comparator. The current study sought to evaluate whether RA may explain differences between MBIs and other *bona fide* treatments in Goldberg et al.'s meta-analysis. This is an important question to resolve for future psychotherapy research and implementation efforts using MBIs. If MBIs are indeed superior to other *bona fide* therapies, this supports their implementation. If the observed superiority of MBIs over specific active controls disappears when accounting for allegiance, this highlights the need to more rigorously consider RA when designing and conducting RCTs of MBIs.

To evaluate the impact of RA in Goldberg et al. (2018), we coded RA in the studies that involved a specific active control or an EBT psychological comparison. We then examined five inter-related research questions:

Question 1) Do MBIs show superiority over other therapies when both specific active controls and EBTs are combined?
Question 2) Are differences smaller for studies using EBT comparators?
Question 3) Does RA moderate differences between MBIs and other therapies?
Question 4) Is EBT status associated with RA?
Question 5) Does the relationship between EBT status and outcome persist when controlling for RA?

Consistent with prior RA studies (Munder et al., 2013), we restricted our analyses to comparisons between psychological interventions; comparisons with non-psychological interventions (e.g., medications) were excluded. Due to an interest in understanding the role of RA in Goldberg et al.'s meta-analysis, only the studies included in the original meta-analysis were used.

Method

The current re-analysis involved all studies included in Goldberg et al.'s (2018) meta-analysis that compared an MBI to another psychosocial intervention that was intended to be therapeutic (i.e., *bona fide* treatment). Goldberg et al. searched four databases (PubMed, PsycInfo, Scopus, Web of Science) using the terms "mindfulness" and "random." MBIs were defined as interventions that had mindfulness meditation as a core component with assignment of home meditation practice. Samples were required to have formal diagnoses or clinically elevated symptoms of a psychiatric condition for which there are evidence-based treatments per APA's Division 12. Additional details regarding eligibility and literature search are included in Goldberg et al.

Goldberg et al. (2018) included 58 studies representing 60 unique comparisons between MBIs and *bona fide* psychological treatments (see Supplemental Materials Table I).[2] The following psychiatric conditions were represented across the studies: depression ($k = 14$), anxiety ($k = 9$), pain ($k = 9$), substance use disorders ($k = 7$), smoking ($k = 5$), weight/eating concerns ($k = 5$), sleep disorders ($k = 4$),

Table I. Researcher allegiance items.

Item	Point value	Description
1	+1	If author advocates for treatment or developed treatment
2	+1	If #1 is true, and authors supervised the therapists, were the therapists in their own condition, or the therapists were extensively trained in the treatment
3	+1	If therapists received more supervision/training than other treatment
4	−1	If supervisor is not a recognized expert in treatment
5	−1	If treatment protocol manual was altered by removing ingredient(s) or changing order in a theoretically deleterious manner
6	−1	If therapists were proscribed from responding in ways a reasonable therapist would routinely do AND proscription was egregious (i.e., proscription was judged to be deleterious to treatment)
7	+1	If greater face-to-face dosage compared with other treatment

Note: Researcher allegiance coding system adopted from Yulish et al. (2017). If item was unclear, it was coded as zero.

posttraumatic stress disorder ($k = 3$), psychotic disorders ($k = 2$).

Effect sizes in Hedges' g units were extracted from Goldberg et al. (2018) reflecting between-group effects (i.e., mindfulness vs. *bona fide* comparison) on targeted symptoms (e.g., symptoms of depression for studies requiring a diagnosis or clinically elevated symptoms of depression for inclusion). A positive effect size reflected greater symptom improvement in the mindfulness condition relative to the control group. Although Goldberg et al. analyzed both post-treatment and follow-up data, only post-treatment effect sizes were analyzed here due to the reduced number of comparisons with follow-up and associated reduction in statistical power for tests of moderation.

RA Coding

RA coding followed the procedure described by Yulish et al. (2017) which was based on those of Munder and colleagues (Munder et al., 2011, 2012, 2013). The seven specific RA items used are shown in Table I. The Introduction and Methods sections of the eligible studies were reviewed for RA coding (i.e., the Results and Discussion were not reviewed). Consistent with prior studies (e.g., Yulish et al., 2017), items for which necessary information was not reported were coded as 0. The allegiance score was calculated by first summing the items pertaining to the MBI or *bona fide* treatment comparison condition. Then, the allegiance score for the *bona fide* treatment comparison was subtracted from the allegiance score for the MBI Equation (1). Thus, a positive value indicated allegiance towards MBIs, zero indicated no allegiance (or equal allegiance), and a negative value indicated allegiance towards the *bona fide* treatment comparison.

$$\Delta\text{Allegiance} = \text{Allegiance}_{\text{MBI}} - \text{Allegiance}_{\text{Bona fide comparison}} \quad (1)$$

RA ratings were done by two psychology professors with expertise in meta-analysis and psychotherapy research. Ratings were made without reference to study results. An intraclass correlation coefficient (*ICC*) was computed using the "icc" function in the "MAd" package (Del Re & Hoyt, 2014). As overall RA was based on the average of all RA items across the two raters, the *ICC* was computed as consistency type with average as the unit of analysis. The initial agreement was good ($ICC = .85$; Cicchetti, 1994). Following initial coding, discrepancies were discussed by the two raters until a consensus was reached.

Due to the expectation that many studies would not report detailed information necessary for coding some RA items, it was decided *a priori* to operationalize RA in two ways. The first involved a raw metric of RA based on the RA discrepancy between the mindfulness and *bona fide* comparison group Equation (1). In addition, a simplified metric was computed by recoding the raw RA ratings into −1 (RA < 0, reflecting RA towards the *bona fide* comparison), 0 (RA = 0, reflecting no or equal RA to both MBI and *bona fide* comparison), and 1 (RA > 0, reflecting RA towards the MBI).

Statistical Analyses

Several distinct analytic approaches were used to address our five research questions. Analyses initially sought to determine the unconditioned omnibus effect size when results were aggregated across what Goldberg et al. (2018) had coded as specific active controls and EBTs (Question 1). A random effects meta-analysis was conducted with effect sizes weighted by the inverse of their variance (Borenstein, Hedges, Higgins, & Rothstein, 2009). Heterogeneity across studies was characterized by the I^2 value, interpreted based on Higgins, Thompson, Deeks, and Altman (2003). A Q test examined whether the degree of heterogeneity was beyond that expected by chance alone, assuming a homogenous distribution of effect sizes. The "rma" function in the "metafor" package in R (Viechtbauer, 2010) was used.

Subsequent analyses tested potential moderators of the omnibus effect also using the "rma" function. An initial model included EBT status as a predictor with EBT status = 1 and non-EBT status = 0 (Question 2). The RA-outcome association was assessed in a subsequent model that included RA as a predictor (Question 3). Separate models examined the raw and simplified versions of RA. A Q test assessed whether EBT status or RA significantly predicted variation in effect sizes across studies.

Next, ordinary least squares (OLS) regression was used to examine EBT status as a predictor of RA (Question 4). OLS is appropriate for this as each comparison included a single rating of EBT status and RA (i.e., they were not nested). Separate models examined the raw and simplified versions of RA using the "lm" function in R (R Core Team, 2018). Standardized regression coefficients were extracted using the "MBESS" package (Kelly, 2007).

Lastly, a final model examined EBT status as a predictor of outcomes when controlling for RA again using the "rma" function in the "metafor" package (Question 5; Viechtbauer, 2010). As for Question 3, separate models examined the raw and simplified versions of RA.

Results

Sixty comparisons drawn from 58 studies were included representing 5,627 participants who were randomly assigned to an MBI or *bona fide* comparison condition. Thirty-seven comparisons conditions were coded by Goldberg et al. (2018) as specific active controls and 23 were coded as EBTs. The mean RA scores were 0.98 ($SD = 1.32$, range = −3–4) and 0.55 ($SD = 0.70$, range = −1–1), for the raw and simplified versions, respectively.

Question 1: Do MBIs Show Superiority Over Other Therapies When Both Specific Active Controls and EBTs Are Combined?

The unconditioned omnibus effect size for MBI versus *bona fide* treatment (Table II, Model 1) was $g = 0.13$, 95% confidence interval (CI) [0.03, 0.23], reflecting a very small effect based on Cohen's (1988) guidelines. Heterogeneity was high ($I^2 = 69.7\%$) and beyond that expected by chance ($Q[59] = 179.29$, $p < .001$).

Question 2: Are Differences Smaller for Studies Using EBT Comparators?

Consistent with Goldberg et al.'s (2018) analyses, EBT status was a significant moderator of between-group effects, with smaller differences between mindfulness and EBT control conditions than non-EBT control conditions ($Q[1] = 6.43$, $p = .011$, Table II, Model 2). EBT status was associated with a reduction in between-group effects ($g = -0.26$ [−0.46, −0.06]) and accounted for 13.0% of heterogeneity in effect sizes.

Question 3: Does RA Moderate Differences Between MBIs and Other Therapies?

Raw and simplified RA ratings both significantly predicted variance in effect sizes (Table II, Models 3 and 4). A one-unit increase in raw RA (i.e., increased allegiance towards MBI) was associated with an increase of $g = 0.10$ [0.02, 0.18], in favour of the MBI (Figure 1).[3] Raw RA accounted for 5.1% of the heterogeneity, although substantial heterogeneity remained even when raw RA was modelled ($I^2 = 68.4\%$, $Q[58] = 171.76$, $p < .001$). Of note, the intercept, which reflects the predicted between-group effect size when RA is equal to zero, was very small and non-significant ($g = 0.02$ [−0.12, 0.15]). A similar pattern was observed for the simplified RA ratings. Any RA towards MBIs (i.e., RA > 0) was associated with an increase in $g = 0.29$ [0.14, 0.44], in favour of the MBI. Simplified RA accounted for 16.8% of the heterogeneity, although again substantial heterogeneity remained even when simplified RA was modelled ($I^2 = 65.6\%$, $Q[58] = 160.41$, $p < .001$). As with the raw RA model, the intercept, which reflects the predicted between-group effect size when RA is equal to zero, was again very small and non-significant ($g = -0.05$ [−0.19, 0.08]).

Table II. Results of meta-analyses examining researcher allegiance-outcome relationship.

Model	*B*	95% *CI*	*p*	*Q*	*df*	*p*(*Q*)	I^2
Model 1 (Unconditioned model)				179.29	59	< .001	69.7%
Intercept	0.13	[0.03, 0.23]	.010				
Model 2				169.23	58	< .001	66.4%
Intercept	0.22	[0.10, 0.34]	< .001				
EBT status	−0.26	[−0.46, −0.06]	.011				
Model 3				171.76	58	< .001	68.4%
Intercept	0.02	[−0.12, 0.15]	.818				
Raw RA	0.10	[0.02, 0.18]	.012				
Model 4				160.41	58	< .001	65.6%
Intercept	−0.05	[−0.19, 0.08]	.441				
Simplified RA	0.29	[0.14, 0.44]	< .001				
Model 5				166.11	57	< .001	66.6%
Intercept	0.11	[−0.06, 0.28]	.189				
EBT status	−0.19	[−0.41, 0.02]	.079				
Raw RA	0.07	[−0.01, 0.15]	.078				
Model 6				156.80	57	< .001	64.3%
Intercept	0.02	[−0.14, 0.19]	.771				
EBT status	−0.15	[−0.36, 0.05]	.142				
Simplified RA	0.25	[0.09, 0.40]	.002				

Note: Raw RA used difference between RA towards mindfulness-based intervention and RA towards *bona fide* treatment comparison condition. Simplified RA coded difference in RA into three-category form (i.e., −1, 0, 1). RA = researcher allegiance; *Q* = *Q* test of heterogeneity.

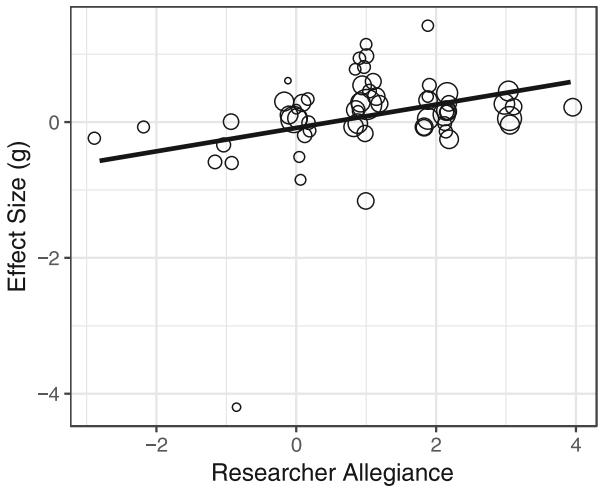

Figure 1. Researcher allegiance towards mindfulness-based interventions is associated with larger between-group effects in favour of mindfulness-based interventions versus *bona fide* treatment comparison conditions. The size of points is relative to their weight in the meta-analysis (i.e., inverse of effect size variance). Figure displays raw researcher allegiance (i.e., not simplified into three-category version). Effect size is Hedges' *g*.

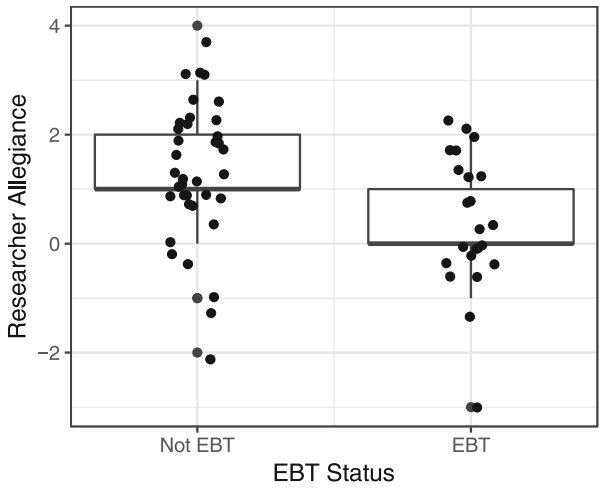

Figure 2. Evidence-based treatment (EBT) status *bona fide* treatment comparison condition is associated with decreased researcher allegiance towards mindfulness-based interventions. Figure displays raw researcher allegiance (i.e., not simplified into three-category version).

Question 4: Is EBT Status Associated with RA?

Next, we explored the possibility that Goldberg et al.'s (2018) specific active control and EBT coding may have been correlated with RA. In separate models, EBT status was a significant predictor of both raw and simplified RA (βs = −0.36 [−0.60, −0.11] and −0.33 [−0.58, −0.08], for raw and simplified RA, respectively; Figure 2). This indicates that RA scores were lower when the *bona fide* treatment comparison was categorized as an EBT.

Question 5: Does the Relationship Between EBT Status and Outcome Persist When Controlling for RA?

Lastly, we modelled both RA and EBT status to determine whether the smaller differences observed when EBT comparisons were used persisted when controlling for RA. EBT status was no longer a significant predictor of between-group effects when controlling for raw or simplified RA (*g*s = −0.19 [−0.41, 0.02], −0.15 [−0.36, 0.05], when controlling for raw and simplified RA, respectively; Table II, Models 5 and 6).[4]

Discussion

Although the importance of considering researcher allegiance RA in the context of testing MBIs has been noted previously (e.g., MacCoon et al., 2012), to our knowledge no previous meta-analysis investigating MBIs has accounted for RA. The current study examined RA as a potential source of bias that might account for the unexpected superiority of MBIs over *bona fide* specific active controls in a recent comprehensive meta-analysis of MBIs for psychiatric conditions (Goldberg et al., 2018).

In contrast to Goldberg et al. (2018) who examined specific active control and EBTs separately, we computed an overall omnibus that aggregated results across all *bona fide* treatment comparisons. This overall omnibus provided a similar conclusion to that reached by Goldberg et al. (2018), indicating that MBIs were slightly superior to *bona fide* treatment comparison conditions, albeit with a very small effect ($g = 0.13$) and substantial heterogeneity ($I^2 = 69.7\%$). We also replicated Goldberg et al.'s result showing that this difference was smaller when EBT comparisons were used, with EBT status accounting for 13.0% of heterogeneity in effect sizes.

Importantly, RA was also a significant moderator of between-group effect sizes. Consistent with prior studies (Munder et al., 2013), a larger between-group effect favouring MBIs was found when RA towards MBIs was present. RA accounted for between 5.1–16.8% of heterogeneity in effect sizes, depending on whether the raw or simplified (i.e., raw RA coded as −1, 0, or 1) forms of RA were used. Both raw and simplified RA models predicted intercept values very close to zero (*g*s = −0.05–0.02), suggesting that differences between MBIs and *bona fide* treatment comparisons are small when RA is absent or balanced.

Lastly, models examined whether the effect of EBT status may be accounted for by RA. We found that RA was indeed predicted by EBT status with moderate effect sizes (βs = −0.36, −0.33, for raw and simplified RA, respectively). And, the impact of EBT status was no longer significant when controlling for RA in multiple predictor meta-regression models. This supports the notion that Goldberg et al.'s (2018) result indicating MBIs were superior to specific active controls (i.e., non-EBT *bona fide* treatment comparisons) and equivalent to EBT *bona fide* treatment comparisons may be better explained by RA. More broadly, this finding supports RA as an important construct and source of bias in psychotherapy research (e.g., that superiority of EBTs may be accounted for by RA; |Munder et al., 2012, 2013).

The current re-analysis is consistent with a large body of psychotherapy research that has examined direct comparisons between *bona fide* treatments generally concluding that treatment differences are close to zero (e.g., Baardseth et al., 2013; Imel et al., 2008; Kivlighan et al., 2015; Wampold et al., 1997), supporting uniform efficacy across various "brands" of psychological intervention (Wampold & Imel, 2015). The current re-analysis does not, however, suggest that MBIs are not effective treatments, rather that they are simply not more effective than other therapies when RA is accounted for. Indeed, comparisons between MBIs and waitlist conditions in Goldberg et al.'s (2018) meta-analysis provides strong evidence that MBIs do produce moderate effects ($d = 0.55$) across a range of psychiatric conditions. As discussed by Wielgosz, Goldberg, Kral, Dunne, and Davidson (2019), effects of MBIs observed across diverse psychiatric conditions may support the implementation of MBIs and the possibility that these treatments target transdiagnostic mechanisms (e.g., repetitive negative thinking). For this reason, MBIs may be promising for addressing comorbidity (e.g., chronic pain and opioid misuse; Garland et al., in press).

Several future directions are important to consider in light of the current findings. Given that MBIs are likely to reduce symptoms on par with other *bona fide* treatments, MBI research could focus on examining when, if ever, MBIs may be preferred to other *bona fide* treatments by patients (e.g., due to lower side effect profile, patient preference) or providers (e.g., due to ease of delivery, group-based delivery format). It would be useful to examine rates of treatment drop-out for MBIs compared with other *bona fide* treatments to assess potential differences in treatment acceptability. As discussed by Hoge, Philip, and Fulwiler (2019), establishing the relative efficacy (Wampold & Imel, 2015) of MBIs may be best served by comparisons between MBIs and established EBTs, as is done when seeking approval for new pharmacological interventions. Such work may be essential for addressing the gap between evidence supporting MBIs for a range of psychiatric conditions and the lack of available insurance reimbursement for delivering MBIs (Hoge et al., 2019).

The current results demonstrating the influence of RA within MBI research should also encourage researchers testing MBIs to embrace the movement towards open science by working to increase the transparency, replicability, and humility of their work (Lilienfeld, 2017; Open Science Collaboration, 2015). Researchers should be incentivized (e.g., by peer-reviewed journals) to clearly report information necessary for coding allegiance (e.g., supervision provided, supervisor credentials, mindfulness instructor background and training). A widespread lack of this kind of information from RCTs of MBIs has been reported across 16 years of MBI research (Goldberg et al., 2017), making these important aspects of RCTs opaque and difficult to evaluate. Researchers with allegiance towards MBIs who are testing MBIs could also consider adopting adversarial collaborators who do not share this allegiance. Adversarial collaborators with allegiance to comparison conditions being tested (e.g., cognitive behavioural therapy, the contextual model of psychotherapy; Arch et al., 2013; MacCoon et al., 2012) are even better.

Several limitations of the current study are important to note. The first and perhaps most important was coding RA based on research reports. While this is the standard practice for RA coding (Munder et al., 2013) and no viable alternative have been established, it relies on relevant information being reported. Many of the included studies did not provide information necessary for coding all RA items. This may have been especially true in the current sample due to the fact that MBI researchers come from a variety of disciplines (e.g., medical subspecialties) that may or may not be sensitive to RA-related factors more commonly discussed in psychotherapy research and psychiatry. Further, as noted by Munder et al. (2012), it is possible that the way in which a research report is written could be influenced by the results of the study. For example, researchers may show greater RA to a treatment arm found to be superior. Previous meta-analyses assessing RA from researchers' prior publications and current research reports (e.g., Berman et al., 1985) has found the RA-outcome association to be similar using either RA assessment procedure.

Our use of heterogeneous psychiatric conditions was both a strength and limitation. While our results may be generalizable to a wider range of

psychiatric conditions, it may also be that the influence of RA varies across psychiatric conditions in ways that were masked in the current analysis. Similarly, we examined the RA-outcome association across a range of MBIs which increases generalizability but may have masked variation. Unfortunately, we did not have sufficient studies within homogeneous psychiatric conditions or specific MBIs for examining these potential sources of variability. Lastly, some models may have been underpowered, leading to Type II error. In particular, the multiple predictor meta-regression examining EBT status when controlling for RA may have lacked power to detect the EBT effect; coefficients were small (Bs = −0.19, −0.15) but in the expected direction.

MBIs are an increasingly popular psychotherapeutic approach that may hold promise for reducing a range of psychiatric symptoms. However, for this work to move forward, it is vital that methodological features, such as RA, be carefully considered and transparently addressed in research reports. The current study provides evidence suggesting that RA operates within MBIs and encourages the use of EBT and other frontline *bona fide* comparison conditions and adversarial collaboration in order to control for this potentially pernicious source of bias. RCTs and meta-analyses that do not consider RA should be interpreted more cautiously.

Supplemental data

Supplemental data for this article can be accessed at https://doi.org/10.1080/10503307.2019.1664783.

Funding

This work was supported by a Varela Award from the Mind & Life Institute to Simon B. Goldberg and the University of Wisconsin–Madison, Office of the Vice Chancellor for Research and Graduate Education with funding from the Wisconsin Alumni Research Foundation.

Notes

[1] Goldberg et al. (2018) did examine differences in dosage as a moderator of treatment effects, a design feature that has been used previously as an indicator of allegiance (Yulish et al., 2017). When examined within comparison type, Goldberg et al. report that dosage matching did not moderate effects. However, Goldberg et al. did not identify this as an aspect of allegiance.

[2] Bowen et al. (2014) and Garland, Roberts-Lewis, Tronnier, Graves, and Kelley (2016) both had two *bona fide* comparison conditions against which an MBI was tested.

[3] Both raw and simplified RA remained significant predictors of outcomes when the negative effect size outlier (Chavooshi, Mohammadkhani, & Dolatshahee, 2016) was excluded (gs = 0.08, 0.23, ps < .05, for raw and simplified RA, respectively).

[4] An anonymous reviewer suggested an additional set of analyses in order to examine the degree to which allegiance may be confounded with aspects of study quality. In particular, the reviewer suggested we predict outcomes with the first allegiance item alone (i.e., whether author advocates for treatment or developed treatment) and whether the remaining allegiance items (i.e., items 2 through 7) mediate the relationship of the first item with outcome. Item 1 and the sum of items 2 through 7 were correlated (r = .46 [0.21, 0.70]). Further, we found that item 1 did predict outcomes in the expected direction (g = 0.20 [0.02, 0.37]) as did the sum of items 2 through 7 (g = 0.11 [0.01, 0.22]). Neither item 1 nor the sum of items 2 through 7 remained significant when entered simultaneously into a meta-regression model (gs = 0.14 [−0.05, 0.34] and 0.08 [−0.04, 0.19], for item 1 and the sum of items 2 through 7, respectively). Thus, it appears possible (based on the non-significant effect of item 1 when controlling for items 2 through 7) that methodologically favouring one treatment may be a pathway through which allegiance is expressed and impacts outcome.

ORCID

Simon B. Goldberg http://orcid.org/0000-0002-6888-0126
Raymond P. Tucker http://orcid.org/0000-0001-9256-6113

References

Arch, J. J., Ayers, C. R., Baker, A., Almklov, E., Dean, D. J., & Craske, M. G. (2013). Randomized clinical trial of adapted mindfulness-based stress reduction versus group cognitive behavioral therapy for heterogeneous anxiety disorders. *Behaviour Research and Therapy*, *51*(4), 185–196.

Baardseth, T. P., Goldberg, S. B., Pace, B. T., Wislocki, A. P., Frost, N. D., Siddiqui, J. R., … Wampold, B. E. (2013). Cognitive-behavioral therapy versus other therapies: Redux. *Clinical Psychology Review*, *33*, 395–405. doi:10.1016/j.cpr.2013.01.004

Berman, J. S., Miller, R. C., & Massman, P. J. (1985). Cognitive therapy versus systematic desensitization: Is one treatment superior? *Psychological Bulletin*, *97*(3), 451–461.

Black, D. S. (2014). Mindfulness-based interventions: An antidote to suffering in the context of substance use, misuse, and addiction. *Substance Use & Misuse*, *49*, 487–491. doi:10.3109/10826084.2014.860749

Bohlmeijer, E., Prenger, R., Taal, E., & Cuijpers, P. (2010). The effects of mindfulness-based stress reduction therapy on mental health of adults with a chronic medical disease: A meta-analysis. *Journal of Psychosomatic Research*, *68*(6), 539–544. doi:10.1016/j.jpsychores.2009.10.005

Borenstein, M., Hedges, L. V., Higgins, J. P. T., & Rothstein, H. R. (2009). *Introduction to meta-analysis*. New York, NY: Wiley.

Bowen, S., Witkiewitz, K., Clifasefi, S. L., Grow, J., Chawla, N., Hsu, S. H., … Larimer, M. E. (2014). Relative efficacy of mindfulness-based relapse prevention, standard relapse prevention, and treatment as usual for substance use disorders: A randomized clinical trial. *JAMA Psychiatry*, *71*(5), 547–556.

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.

Cicchetti, D. (1994). Guidelines, criteria, and rules of thumb for evaluating normed and standardized assessment instruments in psychology. *Psychological Assessment*, *6*(4), 284–290.

Clarke, T. C., Barnes, P. M., Black, L. I., Stussman, B. J., & Nahin, R. L. (2018). *Use of yoga, meditation, and chiropractors among U.S. adults aged 18 and over*. Hyattsville, MD: National Center for Health Statistics.

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed.). Hillsdale, NJ: Erlbaum.

Del Re, A. C., & Hoyt, W. T. (2014). *MAd: Meta-analysis with mean differences*. R package version 0.8-2. Retrieved from http://CRAN.R-project.org/package=MAd

Garland, E. L., Hanley, A. W., Riquino, M. R., Reese, S. E., Baker, A. K., Salas, K., … Howard, M. O. (in press). Mindfulness-oriented recovery enhancement reduces opioid misuse risk via analgesic and positive psychological mechanisms: A randomized controlled trial. *Journal of Consulting and Clinical Psychology*.

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. doi:10.1037/a0035798

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. doi:10.1016/j.brat.2015.11.012

Goldberg, S. B., Tucker, R. P., Greene, P. A., Davidson, R. J., Wampold, B. E., Kearney, D. J., & Simpson, T. L. (2018). Mindfulness-based interventions for psychiatric disorders: A systematic review and meta-analysis. *Clinical Psychology Review*, *59*, 52–60. 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

Head, M. L., Holman, L., Lanfear, R., Kahn, A. T., & Jennions, M. D. (2015). The extent and consequences of p-hacking in science. *PLOS Biology*, *13*(3), e1002106.

Helmes, E., & Ward, B. G. (2017). Mindfulness-based cognitive therapy for anxiety symptoms in older adults in residential care. *Aging & Mental Health*, *21*(3), 272–278.

Higgins, J. P., Thompson, S. G., Deeks, J. J., & Altman, D. G. (2003). Measuring inconsistency in meta-analyses. *BMJ*, *327* (7414), 557–560.

Hoge, E. A., Philip, S. R., & Fulwiler, C. (2019). Considerations for mood and emotion measures in mindfulness-based intervention research. *Current Opinion in Psychology*, *28*, 279–284. doi:10.1016/j.copsyc.2019.02.001

Imel, Z. E., Wampold, B. E., Miller, S. D., & Fleming, R. R. (2008). Distinctions without a difference: Direct comparisons of psychotherapies for alcohol use disorders. *Psychology of Addictive Behaviors*, *22*(4), 533–543. doi:10.1037/a0013171

Kabat-Zinn, J. (1990). *Full catastrophe living: Using the wisdom of your body and mind to face stress, pain, and illness*. New York, NY: Delta.

Kelly, K. (2007). Methods for the behavioral, educational, and social sciences: An R package. *Behavior Research Methods*, *39*(4), 979–984.

Kivlighan, D. M., Goldberg, S. B., Abbas, M., Pace, B. T., Yulish, N. E., Thomas, J. G., … Wampold, B. E. (2015). The enduring effects of psychodynamic treatments vis-a-vis alternative treatments: A multilevel longitudinal meta-analysis. *Clinical Psychology Review*, *40*, 1–14. doi:10.1016/j.cpr.2015.05.003

Kristeller, J. L., & Wolever, R. Q. (2010). Mindfulness-based eating awareness training for treating binge eating disorder: The conceptual foundation. *Eating Disorders*, *19*, 49–61. doi:10.1080/10640266.2011.533605

Langer, Á. I., Cangas, A. J., & Gallego, J. (2010). Mindfulness-based intervention on distressing hallucination-like experiences in a nonclinical sample. *Behaviour Change*, *27*, 176–183.

Leykin, Y., & DeRubeis, R. J. (2009). Allegiance in psychotherapy outcome research: Separating association from bias. *Clinical Psychology: Science and Practice*, *16*(1), 54–65.

Lilienfeld, S. O. (2017). Psychology's replication crisis and the grant culture: Righting the ship. *Perspectives on Psychological Science*, *12*(4), 660–664.

Lorig, K., Holman, H., Sobel, D., Laurent, D., Minor, M., Sobel, D., … Gonzalez, V. M. (2012). *Living a healthy life with chronic conditions: Self-management of heart disease, arthritis, diabetes, asthma, bronchitis, emphysema and other physical and mental health conditions* (4th ed.). Boulder, CO: Bull Publishing Company.

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. doi:10.1016/j.brat.2011.10.011

Maxwell, S. E., Lau, M. Y., & Howard, G. S. (2015). Is psychology suffering from a replication crisis? What does "failure to replicate" really mean? *American Psychologist*, *70*(6), 487–498. doi:10.1037/a0039400

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.

Munder, T., Brütsch, O., Leonhart, R., Gerger, H., & Barth, J. (2013). Researcher allegiance in psychotherapy outcome research: An overview of reviews. *Clinical Psychology Review*, *33*(4), 501–511. doi:10.1016/j.cpr.2013.02.002

Munder, T., Flückiger, C., Gerger, H., Wampold, B. E., & Barth, J. (2012). Is the allegiance effect an epiphenomenon of true efficacy differences between treatments? A meta-analysis. *Journal of Counseling Psychology*, *59*(4), 631–637.

Munder, T., Gerger, H., Trelle, S., & Barth, J. (2011). Testing the allegiance bias hypothesis: A meta-analysis. *Psychotherapy Research*, *21*(6), 670–684.

Open Science Collaboration. (2012). An open, large-scale, collaborative effort to estimate the reproducibility of psychological science. *Perspectives on Psychological Science*, *7*(6), 657–660.

Open Science Collaboration. (2015). Estimating the reproducibility of psychological science. *Science*, *349*(6251), aac4716. doi:10.1126/science.aac4716

R Core Team. (2018). *R: A language and environment for statistical computing*. Vienna: R Foundation for Statistical Computing. Retrieved from https://www.R-project.org/

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., Williams, J. M. G., & Teasdale, J. D. (2013). *Mindfulness-based cognitive therapy for depression* (2nd ed.). New York, NY: Guilford Press.

Society of Clinical Psychology. (2019, June). *Mindfulness-based cognitive therapy*. Retrieved from https://www.div12.org/treatment/mindfulness-based-cognitive-therapy/

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, NY: Routledge.

Wampold, B. E., Mondin, G. W., Moody, M., Stich, F., Benson, K., & Ahn, H. (1997). A meta-analysis of outcome studies comparing bona fide psychotherapies: Empirically, "all must have prizes". *Psychological Bulletin*, *122*, 203–215.

Wieczner, J. (2016, March). Meditation has become a billion-dollar business. *Fortune*. Retrieved from http://fortune.com/2016/03/12/meditation-mindfulness-apps/

Wielgosz, J., Goldberg, S. B., Kral, T. R. A., Dunne, J. D., & Davidson, R. J. (2019). Mindfulness meditation and psychopathology. *Annual Review of Clinical Psychology*, *15*, 1–32. doi:10.1146/annurev-clinpsy-021815- 093423

Yulish, N. E., Goldberg, S. B., Frost, N. D., Abbas, M., Oleen-Junk, N. A., Kring, M., & Wampold, B. E. (2017). The importance of problem-focused treatments: A meta-analysis of anxiety treatments. *Psychotherapy*, *54*(4), 321–338. doi:10.1037/pst0000144