Response to Van Dam and Galante's correspondence "Underestimating harm in Mindfulness-Based Stress Reduction."

Matthew J. Hirshberg, Ph.D.[1], Simon B. Goldberg, Ph.D.[1,2], Melissa Rosenkranz, Ph.D.[1,3], & Richard J. Davidson Ph.D.[1,3,4]

1. Center for Healthy Minds, University of Wisconsin Madison
2. Department of Counseling Psychology, University of Wisconsin Madison
3. Department of Psychiatry, University of Wisconsin Madison
4. Department of Psychology, University of Wisconsin Madison

Word count: 1490

Corresponding author:

Matthew J. Hirshberg
Center for Healthy Minds
University of Wisconsin Madison
625 West Washington Avenue, Madison, WI, 53703
608-262-0035
Email: hirshberg@wisc.edu

Response to Van Dam and Galante's correspondence "Underestimating harm in Mindfulness-Based Stress Reduction."

We thank Van Dam and Galante for their comments. In our view, a misconstrual is foundational to several of their critiques. In their introduction, they construct their critique around the following quote from our paper – "We interpret these data as strong evidence that MBSR is no more harmful than no treatment." They go on to suggest that we ignore other forms of harm, but omitted the key qualifying clause at the end of this sentence; "on the indices of harm we estimated." We intended this clause to acknowledge the limited application of the results (i.e., to indices of harm we measured). In agreement with them, we state that "there are other domains [of harm] in need of investigation."

Van Dam and Galante draw a conclusion from this partial quote that we did not. The sentence in question is an inference based on null hypothesis testing. As one cannot prove the null, in the absence of a significant difference between groups, we can only claim that there is no evidence that one group is more harmful than the other. However, Van Dam and Galante construe our inference as claiming that MBSR "does no harm." We deliberately used language such as "we find no evidence that rates of harm following MBSR are significantly greater" and "these data suggest that MBSR is no more harmful" that reflect the limits of what we could infer. The manuscript title "Prevalence of Harm in Mindfulness-Based Stress Reduction" implies that harm could be present.

Inappropriate conclusion regarding preventative potential of MBSR

Van Dam and Galante raise a number of important concerns in this section, but do not mention that we acknowledged many of them in Limitations. For example, we describe the differences in the community and RCT data and how they limit comparability, and we discuss potential sources of bias leading to overestimation of community MBSR effects. Indeed, we noted that "because of sample differences, these data do not allow us to conclude that MBSR is protective against base rates of symptom increases." Van Dam and Galante also do not mention that we tested two contrasts on all indices of harm: one between community MBSR and RCT wait-list, and another between RCT MBSR and RCT wait-list. We observed no instance in which either community or RCT MBSR led to higher levels of harm than RCT wait-list. Even if we agree that "one cannot reasonably compare an active, open-label intervention to the control arm of an RCT," our primary conclusions remain valid.

Considerable missing data with imputation of low prevalence events

Properly specified multiple imputation (MI) produces unbiased estimates when data are missing at random (MAR; White et al., 2011). Van Dam and Galante make the case that because harm is assumed to be a low prevalence event and we described a limited number of predictors in the imputation model, MAR assumptions are too strong. However, our MI models included nearly all available information: gender, participation year, participant age, data type (i.e., community MBSR, RCT MBSR or RCT wait-list), and pre- and post-test score on the outcome. Of the roughly 23% of community MBSR data missing at post-test, dropout rates were very low (2%). Explanations for missing data that seem to us as plausible as adverse events include that survey completion was optional, had no impact on enrollment or participation, was not part of

research, and did not involve compensation. In addition, baseline symptom severity was not related to missingness.

Taking their concern seriously, we quantitatively evaluate the robustness of our results to missing not at random (MNAR) assumptions. Even under the most implausible missingness scenario (i.e., all missing data represent harm), harm in community MBSR remains significantly less likely than RCT wait-list on multiple indices of psychological and physical symptoms (average change, proportion with worsening symptoms, proportion with >35% increase in symptoms). In order for clinically significant harm on the GSI to be significantly more likely in community MBSR than in RCT wait-list, our estimates would have to be underestimated by nearly 450%. That is, more than half of the participants missing post-test data would need to have experienced clinically significant harm. If we assume missed harm events are as likely in the missing RCT data as in the community data, underestimation of harm in community MBSR would need to be even greater to invalidate our findings.

Limited measurement of harm

We agree that following-up with participants after MBSR is important and post-MBSR harm events are possible. In our view, attributing those potential harms to MBSR is a logical fallacy. Rather, this would indicate that practicing meditation after MBSR may lead to harm. Akin to meditation retreats after MBSR, imagine an individual participates in a running camp to train for a 10-kilometer race, runs the race without injury, and then a month after the camp ends, decides to run a marathon and gets injured. To claim that the injury is the result of the running camp seems specious. It also may preclude understanding the proximal causes of the injury (e.g., intensive retreat practice).

Potential sample bias

We generally agree with Van Dam and Galante's points. We add to them that some of the "inflated effect size" of behavioral interventions such as MBSR outside of controlled trials may be due to the fact that motivation and engagement are central to behavioral change (influencing expectancy and other non-specific factors; Wampold & Imel, 2015). RCTs may enroll participants whose levels of motivation are lower than those who self-select into community treatment.

Collation of data across potentially different trials

Van Dam and Galante raise two putative issues here: collation of data across trials and lack of transparency. On the latter, we agree that transparency is important. While we have widely adopted open science practices – our group routinely pre-registers studies and shares code and data – we feel the utility of "prospective" registration of hypothesis testing on extant data, especially when data have been previously processed, analyzed, and published (i.e., these RCT data) is questionable. Additionally, RCTs 1 and 2 were conducted before pre-registration was common. We note that all outcomes collected by the community MBSR program were reported. Links to RCT registrations are provided below.

It is true that variation across trials may be a source of variation in the data. In our view, the benefits to statistical power warranted collation. Moreover, the trials were similar in important ways: conducted by the same research group, in the same geographic location, using similar inclusion/exclusion criteria, with MBSR conducted by the same community MBSR program and community MBSR teachers over a roughly contemporaneous period as the community MBSR data.

Inappropriate conclusion regarding preventative potential of MBSR

As previously described, we do not claim that MBSR is preventative of harm. Further, we did not, as Van Dam and Galante state, "use the deterioration of the wait-list group, compared to the improvement of the community group as a means to saying that community MBSR is preventative of harm." Justification for the claim that MBSR *may* be preventative of harm relied on analyses in which we compared rates of harm in MBSR and in wait-list (e.g., proportion with >35% symptom worsening). As noted, comparisons were also made between RCT MBSR and RCT wait-list.

Van Dam and Galante also invoke the possibility that increased symptoms in the wait-list were due to the "nocebo effect," citing Furukawa et al's (2014) network meta-analysis. While theoretical claims have been made in support of nocebo effects (e.g., resentful demoralization; Heppner et al., 2016), as discussed by Munder et al. (2019), there is little empirical evidence that this is the case. Munder et al. highlight important limitations specifically related to Furukawa et al.'s paper, namely that none of the no treatment comparisons (to which wait-list was indirectly compared) included patients seeking treatment. This sample difference produced small effects for the no treatment comparisons, resulting in the inaccurate impression that no treatment was more effective than wait-list.

Van Dam and Galante appear to have interpreted "strong evidence against claims that MBSR may increase harm" as stating that "MBSR does no harm." Although we remain "circumspect about the evidence that MBSR is protective," we disagree with the possibility that the evidence for lower rates of harm in MBSR was driven by wait-list nocebo effects. We also believe a wait-list control is more appropriate than an active control comparison; active control

comparisons would not address whether rates of harm in MBSR are different than population-level base rates of harm.

Conclusions

Like Van Dam and Galante, we believe that understanding the potential for harm in meditation interventions is important. While our paper provides the first quantitative estimates of multiple indices of harm across multiple harm domains in a large sample of MBSR participants, we agree that more research is needed. We hope that the points raised by Van Dam and Galante along with our responses support future work in this area and aid in informing clinicians and the public about the potential risks and benefits of MBSR.

RCT Registrations:
https://clinicaltrials.gov/ct2/show/NCT01301105
https://clinicaltrials.gov/ct2/show/NCT01057368
https://clinicaltrials.gov/ct2/show/study/NCT02157766

Financial support

Financial support was provided by a National Academy of Education / Spencer Foundation postdoctoral research fellowship to MJH, the National Institutes of Health (U01AT002114-01A1 and P01AT004952) to RJD, MK and Antoine Lutz, the National Center For Complementary & Integrative Health (K23AT010879) to SBG, and by generous donations to the Center for Healthy Minds. The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health.

Conflicts of interest

Matthew J. Hirshberg is a contracted provider at the community MBSR providing clinic. Richard J. Davidson is the founder, president, and serves on the board of directors for the non-profit organization, Healthy Minds Innovations, Inc. In addition, RJD served on the board of directors for the Mind & Life Institute from 1992-2017. No donors, either anonymous or

identified, have participated in the design, conduct, or reporting of research results in this manuscript.

Statement of Ethics

The authors assert that all procedures contributing to this work comply with the ethical standards of the relevant national and institutional committees on human experimentation and with the Helsinki Declaration of 1975, as revised in 2008. All participants in all of the studies included in this manuscript provided their written, informed consent before participating. All methods and procedures were reviewed and approved by the University of Wisconsin Madison Institutional Review Board.

Author Contributions

Matthew J. Hirshberg led sensitivity analyses and writing. All authors participated in the writing of the rejoinder. All authors have provided final approval of the manuscript for submission.

References

Furukawa, T. A., Noma, H., Caldwell, D. M., Honyashiki, M., Shinohara, K., Imai, H., Chen, P., Hunot, V., & Churchill, R. (2014). Waiting list may be a nocebo condition in psychotherapy trials: A contribution from network meta-analysis. *Acta Psychiatrica Scandinavica*, *130*(3), 181–192. https://doi.org/10.1111/acps.12275

Heppner, P. P., Wampold, B. E., Owen, J., Thompson, M. N., & Wang, K. (2016). Research Design in Counseling (4th Ed.). Boston: Cengage.

Lindahl, J. R., Fisher, N. E., Cooper, D. J., Rosen, R. K., & Britton, W. B. (2017). The varieties of contemplative experience: A mixed-methods study of meditation-related challenges in

Western Buddhists. *PLOS ONE*, *12*(5), e0176239. https://doi.org/10.1371/journal.pone.0176239

Munder, T., Flückiger, C., Leichsenring, F., Abbass, A. A., Hilsenroth, M. J., Luyten, P., ... & Wampold, B. E. (2019). Is psychotherapy effective? A re-analysis of treatments for depression. Epidemiology and Psychiatric Sciences, 28(3), 268-274.

Wampold, B. E., & Imel, Z. E. (2015). *The great psychotherapy debate: The evidence for what makes psychotherapy work*. Routledge.

White, I. R., Royston, P., & Wood, A. M. (2011). Multiple imputation using chained equations: Issues and guidance for practice. *Statistics in Medicine*, *30*(4), 377–399. https://doi.org/10.1002/sim.4067