

# Alternative meta-analysis of behavioral interventions to promote action on climate change yields different conclusions

Sander van der Linden<sup>1✉</sup> & Matthew H. Goldberg<sup>2</sup>

ARISING FROM C. F. Nisa et al. *Nature Communications* <https://doi.org/10.1038/s41467-019-12457-2> (2019)

The degree to which behavioral interventions can promote widespread action on climate change is an important and urgent question. In their recent meta-analysis, Nisa et al.<sup>1</sup> conclude that behavioral interventions have alarmingly small effects. Here we re-analyze the data and conclude that the meta-analytic effect reported by Nisa et al.<sup>1</sup> is highly sensitive to the chosen estimator and, across a range of alternative estimators, we find that the average effect-size of behavioral interventions is at least twice as large as initially reported. These results are important not only because they impact estimates of the potential for population-level behavior change but also because they have implications for how behavioral researchers should conduct meta-analyses.

We applaud Nisa et al.<sup>1</sup> for their important meta-analysis on randomized controlled trials, which evaluated behavioral interventions to promote household action on climate change. We agree with the authors that no clear consensus exists on the question of which behavioral interventions are most effective in promoting individual and collective action on climate change. Yet, we were surprised and concerned to learn that across the 83 interventions published between 1976 and 2017 that met the inclusion criteria—which jointly included over 3 million individuals—the average standardized effect size across all behavioral interventions was only  $d = -0.093$  (95% CI;  $-0.16, -0.06$ ). This effect size is small with one side of the confidence interval relatively close to zero. The authors translate this effect into a probability of benefit (i.e., probability that the intervention will promote climate change mitigation behaviors in the experimental vs. control group) of 6.6%. The authors conclude: “behavioral interventions have a very small positive effect” and “our work indicates alarmingly low levels of behavioral plasticity” (p.7). Because this is a strong conclusion with policy-implications, it is important that the data support this claim.

Here we make the straightforward point that unfortunately the data do not support the conclusions reported by Nisa et al.<sup>1</sup>. In fact, the reader might be interested to learn that the results of the

meta-analysis are highly sensitive to the estimation method. For example, the level of heterogeneity in the reported meta-analysis is very high (as measured by the  $I^2$  statistic), up to 98% for some interventions (e.g., transportation) with an overall average of about 65%. Nisa et al.<sup>1</sup> estimated a random-effects meta-analysis with the DerSimonian–Laird (DL) procedure, a common method for modeling between-study variance. However, when heterogeneity is moderate to high—as is the case in the current study—this estimator is biased and not recommended<sup>2</sup>. Although the DL method is relatively simple and popular, it can lead to severe underestimation of the variance when either the number of studies is limited or the heterogeneity is large<sup>3</sup>. Instead, Empirical Bayes or Restricted Maximum Likelihood is often recommended<sup>2</sup>, especially when heterogeneity is relatively high.

When we re-estimate Nisa et al.’s<sup>1</sup> meta-analysis using the data provided with Empirical Bayes we get a very different estimate for the average standardized effect size:  $d = -0.204$  (95% CI;  $-0.250, -0.158$ ). Importantly, the estimation method need not be limited to Empirical Bayes to make this point. For example, using other often recommended heterogeneity estimators<sup>4</sup>, such as restricted maximum likelihood (REML), we obtain a very similar estimate of  $d = -0.19$  (95% CI;  $-0.24, -0.15$ ). The Sidik–Jonkan procedure gives an estimate of  $d = -0.22$  (95% CI;  $-0.27, -0.17$ ) and Hedges yields a near-identical estimate of  $d = -0.21$  (95% CI;  $-0.25, -0.16$ ). Under the range of effect-sizes observed here, the probability of benefit (POB) more than doubles—from 6.6% to 14.4%.

This finding should therefore lead to different conclusions about the average efficacy of behavioral interventions. Indeed, based on these results, the conclusion of the paper should probably not be that behavioral interventions have very small effects or alarmingly low levels of behavioral plasticity. It is important to contextualize this difference, because effect sizes in psychology are known to be small<sup>5</sup> and, as others have noted<sup>6</sup>, their effects can accumulate over time with repetition. Following Furukawa and Leucht<sup>7</sup>, we convert Cohen’s  $d$  into the number-needed-to-treat (NNT) using a default control group event rate of

<sup>1</sup>Department of Psychology, School of Biological Sciences, University of Cambridge, Cambridge CB2 1TN, UK. <sup>2</sup>Yale Program on Climate Change Communication, Yale University, New Haven, CT 06520, USA. ✉email: [sander.vanderlinden@psychol.cam.ac.uk](mailto:sander.vanderlinden@psychol.cam.ac.uk)

20%. A  $d$  of 0.09 could be considered trivial and requires, on average, about 38 people to be exposed to the treatment to engender one extra unit of behavior change in the treatment group. Importantly, this number falls by more than 50% to only 16 when  $d = 0.20$ , which is generally considered a meaningful effect size in psychology<sup>5</sup>. In other words, we want to point out to the reader that the main results and overall conclusions of the paper are highly sensitive to the estimation method.

It would have been helpful if the authors had made a case for why their estimation method was appropriate in light of the significant degree of heterogeneity across interventions. Moreover, instead of relying on a single-estimator, the authors could have conducted robustness checks with multiple estimators to provide transparency for the reader around how the estimation method influences the results of the meta-analysis. Such sensitivity analyses are generally recommended as reliance on a single estimator in meta-analyses may lead to non-robust conclusions<sup>4</sup>. For example, as we show here, alternative estimators provide the reader with a more optimistic interpretation about the average efficacy and plasticity of behavioral interventions to promote household action on climate change. Although the average effects of behavioral interventions are still not particularly large in absolute terms ( $d = 0.20$ ), they are not alarmingly small and could be consequential when scaled at population level<sup>5</sup>. At minimum, they are significantly larger than what was reported by Nisa et al.<sup>1</sup>.

### Data availability

The data necessary to reproduce the analyses reported in this paper can be retrieved from: [https://figshare.com/articles/Dataset\\_Nisa\\_et\\_al\\_2019\\_NComms\\_xlsx/9641999](https://figshare.com/articles/Dataset_Nisa_et_al_2019_NComms_xlsx/9641999).

Received: 28 January 2020; Accepted: 7 July 2020;

Published online: 05 August 2020

### References

- Nisa, C. F. et al. Meta-analysis of randomised controlled trials testing behavioural interventions to promote household action on climate change. *Nat. Commun.* **10**, 4545 (2019).
- Sidik, K. & Jonkman, J. N. A comparison of heterogeneity variance estimators in combining results of studies. *Stat. Med.* **26**, 1964–1981 (2007).
- Guolo, A. & Varin, C. Random-effects meta-analysis: the number of studies matters. *Stat. Methods Med. Res.* **26**, 1500–1518 (2017).
- Langan, D., Higgins, J. P. & Simmonds, M. An empirical comparison of heterogeneity variance estimators in 12,894 meta-analyses. *Res. Synth. Methods* **6**, 195–205 (2015).
- Funder, D. C. & Ozer, D. J. Evaluating effect size in psychological research: Sense and nonsense. *Adv. Methods Pract. Psychol. Sci.* **2**, 156–168 (2019).
- Stern, P. C. A reexamination on how behavioral interventions can promote household action to limit climate change. *Nat. Commun.* **11**, 918 (2020).
- Furukawa, T. A., & Leucht, S. How to obtain NNT from Cohen's  $d$ : comparison of two methods. *PLoS One.* **6**, e19070 (2011).

### Author contributions

S.vd.L and M.H.G. conceptualized the research. S.vd.L. and M.H.G. analyzed the data. S.vd.L. drafted the manuscript with input and revisions from M.H.G.

### Competing interests

The authors declare no competing interests.

### Additional information

**Correspondence** and requests for materials should be addressed to S.vd.L.

**Peer review information** *Nature Communications* thanks Julian Higgins for the contribution to the peer review of this work.

**Reprints and permission information** is available at <http://www.nature.com/reprints>

**Publisher's note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



**Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2020