



# The Yet Underestimated Importance of Communicating Findings from Educational Trials to Teachers, Schools, School Authorities, or Policy Makers (Comment on Brady et al. (2023))

Steffen Zitzmann<sup>1</sup> · Nils Machts<sup>2</sup> · Nicolas Hübner<sup>1</sup> · Stefan Schaubert<sup>3</sup> · Jens Möller<sup>2</sup> · Christoph Lindner<sup>4</sup>

Accepted: 3 May 2023 / Published online: 15 June 2023  
© The Author(s) 2023

## Abstract

This article is a comment on Brady et al. (*Educational Psychology Review*, 35, 6–37, 2023) with which we largely agree. We add to this important discussion by pointing to the underestimated importance of communicating findings to stakeholders, which is important because recommendations are derived from them, and a correct understanding is essential for stakeholders to evaluate the usefulness of recommendations and strengthen stakeholders' beliefs in their added value. Moreover, we suggest that in addition to an effect size, responder rates may be communicated so that stakeholders can better understand the consequences of implementing a treatment.

**Keywords** Experimental research · Causality · Science communication · Responder · Responder analysis

Recently, Brady et al. (2023) alerted the field of educational psychology that it is further moving away from experimental research. At the same time, articles with findings from non-experimental work tend to include more and more far-reaching conclusions and recommendations for practice, though these recommendations may not sufficiently be backed up by robust findings about the causal roles of the investigated variables. Therefore, Brady et al. (2023) suggested that researchers should

---

✉ Steffen Zitzmann  
steffen.zitzmann@uni-tuebingen.de

<sup>1</sup> Hector Research Institute of Education Sciences and Psychology, University of Tübingen, Tübingen, Germany

<sup>2</sup> Kiel University, Kiel, Germany

<sup>3</sup> University of Oslo, Oslo, Norway

<sup>4</sup> Max Planck Institute of Psychiatry, Munich, Germany

base recommendations on experimental rather than on non-experimental research. Although we largely agree with the authors, sticking with the blurry dichotomy of “scientific experimental” versus “unscientific non-experimental” work likely falls short. This can be seen from great challenges to replicate findings from randomized experiments, which constitutes the central foundation of “scientific” experimental research (OpenScience, 2015). We strongly agree with Brady et al. (2023) that causality should be at the core of research in educational psychology, particularly if it intends to inform practice. However, whether or not results represent causal effects can only be answered with great knowledge and assumptions about the data, design, and the methods used, and this holds for non-experimental as well as for experimental work (Wadhwa et al., 2019). For a discussion of the various assumptions needed to interpret results causally, see, for example, Hübner et al. (2023) in this journal (see also Grosz et al., 2020).

With this comment, however, we do not want to discuss the justifiability of causal claims but point to the yet underestimated importance of how findings from research in educational psychology are communicated to, for example, teachers, schools, school authorities, or policy makers. This is important because recommendations are derived from the findings, and a correct understanding of the findings will help stakeholders evaluate their usefulness and strengthen their beliefs of whether the recommendations may be worth implementing. Also, describing findings in a meaningful and understandable way is in line with the general movement in educational psychology to increase the outreach of educational psychological research. Any practical recommendations should be preceded by a summary of findings, which should *be meaningful to stakeholders and be communicated in such a way that these persons will easily understand them*.

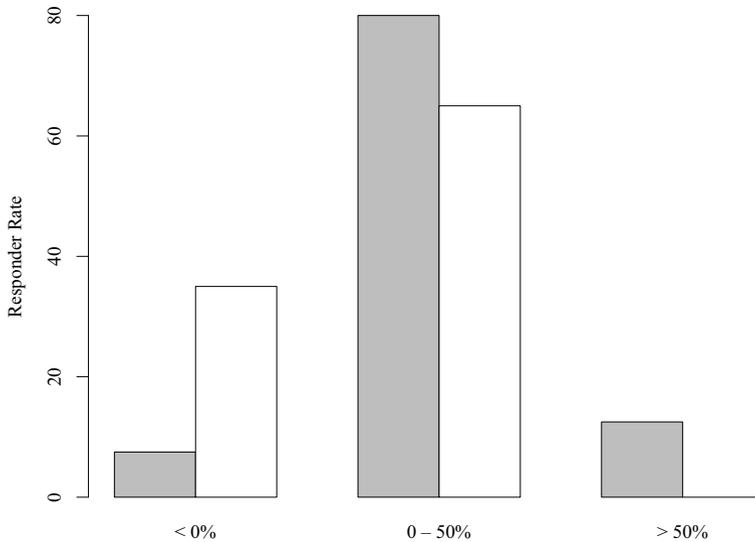
In order to be able to communicate findings in an understandable way, it is necessary to consider the professional background of the stakeholders to whom we report these findings. From the perspective of a teacher, a statistic that is computed as an average across many students, such as an effect size (e.g., a Cohen’s  $d$  type of measure), yields only one value, and there might not be a single student who exhibited exactly this value. Thus, an effect size may be perceived as telling little about the changes of real students (see, e.g., Schmidt et al., *in press*, for teacher perceptions of effect size measures). In a nutshell, successfully communicating effect size measures to stakeholders remains a challenging task despite the existence of some guidelines, such as those from What Works Clearinghouse (WWC).

In addition to an effect size measure, what is needed is a strategy that capitalizes more on individual students. Such a strategy would go well in line with teachers’ everyday observation that, for example, students learn differently. It is interesting to consider what we can learn from other sciences in this regard. In a way, the challenge of communicating findings in educational psychology to stakeholders parallels the situation in clinical research, where stakeholders are physicians and clinical psychologists and where useful solutions for similar challenges have emerged. Although a physician may not be an expert in statistics, they may have a good sense of what an individual patient’s symptom reduction means as well as the capacity to intuitively understand frequencies of patients with comparable symptom reductions. Therefore, in addition to an effect size measure, findings from a so-called *responder analysis* are typically communicated. This type of analysis involves reporting the relative frequencies of

patients whose symptoms reduced. By responder we mean someone who responded in the sense that they improved substantially, without referring to a specific cause for the improvement such as the treatment, the placebo, or the tendency of the patient to improve over time due to spontaneous remission. This lack of a reference to a specific cause is reasonable because in the typical trials involving a treatment and a control group, we cannot determine how, for example, a patient in the treatment group would have improved if this patient had been in the control group (e.g., if they had received a placebo or had not been treated at all). Note that the concept of a response should not be confused with treatment adherence (Schmidt et al., 2012).

The responder analysis can easily be adapted to educational psychology, and not only research on learning but also research on other phenomena, such as effects of intervention programs to reduce stress in students (Regehr et al., 2013), could increase its impact by communicating responder rates. Referring to a student's values at the beginning and end as *pre* and *post*, respectively, their percentage change can be computed as  $(\text{post} - \text{pre}) / \text{pre} \cdot 100\%$  (see, e.g., Leucht et al., 2009; see also Zitzmann et al., *in press*). After the percentages change have been obtained for all students, they can be classified into categories of change based on reasonable cutoffs. It could be criticized that it is not clear how these cutoffs should be chosen and are thus arbitrary. However, the cutoffs may come from theory; for example, from expert considerations of what should at least be learned during a course/period of time. Or, they may come from empirical research (e.g., what can on the basis of previous research be expected to be learned). Another possible critique is related to the problem of classifying students when percentage change is assessed with low reliability (but see Zitzmann et al., 2023, for a way to deal with this problem). Once the students are classified, counting them per category, dividing by the sample size, and multiplying by 100% yield the responder rates, which can be packed into easy to interpret bar charts, such as the one in Fig. 1. To generate the chart, we used cutoffs of 0% and 50% as illustrative, certainly not very valid examples of cutoffs in educational psychological research. However, when theory is yet vague and empirical findings are rare, they may provide a starting point but not a definite definition. The chart presents the responder rates of two different groups, the treatment and the control groups. As can immediately be seen from the figure, the responder rates differed greatly between these groups, with a clear tendency for the treatment group to show higher responder rates in categories of positive gain (i.e., in the 0 to 50% and > 50% categories), whereas this pattern was reversed in the category of no gain (i.e., in the < 0% category). This means that a much larger amount of students gained substantially in the treatment relative to the control group — an interpretation that may help stakeholders evaluate if findings are of practical significance. It is important to note that it is also possible to present responder rates for different subgroups. Such a presentation would better reflect that most effects are heterogeneous (Bryan et al., 2021) and would allow stakeholders to judge if findings are of practical significance to their specific target group.

Taken together, the goal was to comment on Brady et al. (2023). We appreciated their contribution and largely agreed with the authors' suggestion that educational psychologists should base practical recommendations on causal effects. We added to the discussion by pointing to the underestimated importance of communicating findings to stakeholders, which is important as recommendations are derived from the



**Fig. 1** Responder rates from a hypothetical trial. Gray bar = treatment group; white bar = control group

findings, and a correct understanding of them is essential for stakeholders to evaluate the usefulness of recommendations and strengthen stakeholders' beliefs in their added value. We argued that in addition to an effect size measure, reporting of yet other measures can help translate research to practice. Specifically, we suggested that responder rates could be communicated so that stakeholders can better understand the consequences of implementing a treatment in terms of students' learning gains, reductions of stress, or changes in other important student outcomes. As responder rates emerge from individual students' percentages change and are thus an intuitive way of summarizing findings, stakeholders can better infer their practical significance (see Krammer et al., *in press*, for a similar argument). It goes without saying that the efficacy of communicating responder rates needs to be tested empirically *before* conclusive recommendations can be formulated. Such a study may involve testing our assumptions that teachers, schools, school authorities, and policy makers would understand individual students' changes and frequencies of students with comparable changes more easily as well as testing the comprehensibility of the responder analysis as a whole.

**Funding** Open Access funding enabled and organized by Projekt DEAL.

**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 licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence 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 licence, visit <http://creativecommons.org/licenses/by/4.0/>.

## References

- Brady, A. C., Griffin, M. M., Lewis, A. R., Fong, C. J., & Robinson, D. H. (2023). How scientific is educational psychology research? The increasing trend of squeezing causality and recommendations from non-intervention studies. *Educational Psychology Review*, 35, 36–37. <https://doi.org/10.1007/s10648-023-09759-9>
- Bryan, C. J., Tipton, E., & Yeager, D. S. (2021). Behavioural science is unlikely to change the world without a heterogeneity revolution. *Nature Human Behaviour*, 5, 980–989. <https://doi.org/10.1038/s41562-021-01143-3>
- Grosz, M. P., Rohrer, J. M., & Thoemmes, F. (2020). The taboo against explicit causal inference in non-experimental psychology. *Perspectives on Psychological Science*, 15, 1243–1255. <https://doi.org/10.1177/1745691620921521>
- Hübner, N., Wagner, W., Zitzmann, S., & Nagengast, B. (2023). How strong is the evidence for a causal reciprocal effect? Contrasting traditional and new methods to investigate the reciprocal effects model of self-concept and achievement. *Educational Psychology Review*, 35, Advance online publication. <https://doi.org/10.1007/s10648-023-09724-6>
- Krammer, G. (in press). A cautionary note on aggregation in educational psychology and beyond. *Theory & Psychology*. <https://doi.org/10.1177/09593543231172495>
- Leucht, S., Davis, J. M., Engel, R. R., Kissling, W., & Kane, J. M. (2009). Definitions of response and remission in schizophrenia Recommendations for their use and their presentation. *Acta Psychiatrica Scandinavica*, 119, 7–14. <https://doi.org/10.1111/j.1600-0447.2008.01308.x>
- Open Science Collaboration. (2015). Estimating the reproducibility of psychological science. *Science*, 349, 1–8. <https://doi.org/10.1126/science.aac4716>
- Regehr, C., Glancy, D., & Pitts, A. (2013). Interventions to reduce stress in university students A review and meta-analysis. *Journal of affective disorders*, 148, 1–11. <https://doi.org/10.1016/j.jad.2012.11.026>
- Schmidt, K., Edelsbrunner, P. A., Rosman, T., Cramer, C., & Merk, S. (in press). When perceived informativity is not enough. How teachers perceive and interpret statistical results of educational research. *Teaching and Teacher Education*
- Taylor, S., Abramowitz, J. S., & McKay, D. (2012). Non-adherence and non-response in the treatment of anxiety disorders. *Journal of Anxiety Disorders*, 26, 583–589. <https://doi.org/10.1016/j.janxdis.2012.02.010>
- Wadhwa, M., & Cook, T. D. (2019). The set of assumptions randomized control trials make and their implications for the role of such experiments in evidence-based child and ado-lescent development research. *New Directions for Child and Adolescent Development*, 2019, 17–37. <https://doi.org/10.1002/cad.20313>
- Zitzmann, S. (in press). Einzelfallbezogene Veränderungsdiagnostik [Diagnostics of individual change]. In R. Dohrenbusch (Ed.), *Psychologische Begutachtung*. Wiesbaden, Springer
- Zitzmann, S., Lindner, C., Leucht, C., & Leucht, S. (2023). Taking uncertainty in the assessment of response into account An advanced guideline for computing responder rates in clinical trials. *Manuscript submitted for publication*.

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