

Response to “When research ethics is downplayed”



Sophie Havighurst

Mindful Centre for Training and Research in Developmental Health, Department of Psychiatry, The University of Melbourne
sshavi@unimelb.edu.au

Ole Andre Solbakken

Department of Psychology, University of Oslo

Maud Edvoll

Department of Psychology, University of Oslo

Rune Flaaten Bjørk's claims of research misconduct, selective reporting and of 'overselling' findings are not substantiated.



Havighurst, Solbakken og Edvoll

To the Editors,

Thank you for providing us with an opportunity to respond to Bjørk's letter (2025) to your journal. This study and the publication in question have undergone extensive scrutiny and review by experts in the field, and an independent expert committee, all of whom disagree with the conclusions drawn by Bjørk. Presented below is an itemised response to the criticisms made by Bjørk.

Dette er et tilsvarende til artikkelen Når forskningsetikken nedprioriteres.

Faulty role description

It is incorrect that Bjørk was responsible for the implementation of the study. Rather, he was a PhD student working on the project under a Norwegian Principal Investigator (PI). The PI on the project was responsible for and oversaw all aspects of the project, engaged the majority of the group leaders, supervised most of the students, and wrote up the study for publication.

Extensive efforts were made



Contrary to Bjørk's account, many months were spent writing emails and responding to his questions. His suggestions were always considered. Contrary to what he has stated in his criticism, many efforts were made to retain him as a study author. He said he did not wish to be an author because he disagreed with what the paper reported.

Criticism of the work

Bjørk raised his concerns with Espen Røysamb, Head of Research, Department of Psychology, and Bjørn Lau, Head of the Department of Psychology, UiO, he was assured that the analyses appeared appropriate. However, in the name of transparency, they asked UiO's Science Ombud for a recommendation for how to respond. The Science Ombud recommended appointing an independent review committee to closely examine Bjørk's concerns and conduct a fine-grained review of the study method, data, syntax, analyses, results and the final paper itself. The independent review committee found that the methods used were robust and appropriate. It made some recommendations for ways to strengthen the paper surrounding the moderation analyses being speculative and exploratory. Minor adjustments to the wording were made, and a final version of the paper was reviewed again by the committee, which now found it acceptable for publication. The final conclusions of the department regarding this process, as documented by Lau, were as follows:

'We at PSI consider the article, as submitted, to be approved and in line with ethical and research integrity standards. This conclusion follows the recommendation of the independent review committee, which we as PSI leadership support.'

Moderator analyses are appropriate

The final version of the paper approved by the review committee stated explicitly that some moderator analyses (for instance, those exploring the outcome of anxiety in children who were shy) were exploratory. Scientific practice does indeed allow for additional subgroup analyses, even when there are no main effects for the total sample (Deaton & Cartwright, 2018), if they are labelled as exploratory. From the beginning of the study, the intention was to examine the role of parent functioning and child temperament as moderators of the intervention outcomes – and this is precisely what was done in the paper. The PI learned about the process of pre-registration late in the study implementation and registered the study after it had been conducted – which is acceptable practice by the Clinical Trials Registry. This is partly because registering intervention trials has only recently become normal practice, and many intervention trials have not previously been registered because it was not yet common practice to do so.

P-hacked or phished data

The accusation of 'p-hacking' is untenable, since the authors openly distinguish between pre-registered primary hypotheses and these additional explorations. The main analyses outlined in the pre-registration were included in the paper. Additional questions are common to explore that may not be pre-registered because it is usual to register only the main outcomes being explored in clinical

trials registries. Studies like this one will often have different papers that use the data to examine different additional questions, such as moderators and sub-group analyses. In this case, the main effects and exploratory moderation effects (looking at temperament and parent emotion functioning as moderators) were undertaken.



Allegations about the program

This study was not a clinical trial, but a universal preventive intervention aimed at families in the general population. Small yet significant effect sizes are common and meaningful in this type of preventive work. In contrast, a clinical trial would be more likely to result in changes with large effect sizes. Previous Tuning in to Kids (TIK) studies with clinical samples have a more comprehensive intervention. For example, one TIK study with children with emerging conduct disorder included an eight-session parenting programme, an eight-session child programme, a teacher intervention and wider family and school supports (Havighurst et al., 2015). The current N-TIK study is not the same, so using this data to draw conclusions for a clinical population is incorrect. The N-TIK study was for anyone interested in attending a parenting programme with a child in the final year of kindergarten – what is typically called a universal sample. Prevention trials are light-dose interventions and expect small changes because many people have no problems. For this reason, moderator analyses are important because for some sub-groups of the population an intervention might be more effective – and this was what was explored in the current paper. It was explicitly stated that the intervention did not yield a main effect for anxiety with the entire sample, but found some improvement in shyer children, presented as an exploratory finding in the paper. This is not a ‘marketing ploy’ as suggested by Bjørk, but rather a nuanced point that may guide future research. There has been an increasing call for prevention researchers to examine moderators because universal prevention studies where only main effects are examined for the whole sample can mask the fact that for some sub-groups (such as children who are shyer) the intervention does have an effect. This type of analysis has been called for by researchers, such as Jay Belsky, as an important way forward for understanding intervention outcomes (Belsky and van Ijzendoorn, 2015).

Claims of data cherry picking

The paper clarifies why additional data (especially from teachers) was not used in the analyses: it was because of unfortunate administrative errors that resulted in 70% of the teacher data not being collected, with the remaining data being often incorrectly linked across the two time points (matching the wrong children over time or the wrong teacher over time). The PI dismissed the use of any of the data early on, but not in an attempt to conceal unfavourable data – it was simply because the data was unusable and incorrect.

With regard to including only one (not both) of the parents in the study outcomes, it is standard practice to only evaluate with the parent who participated – not the non-participating parent. This is because an intervention does not usually generalise to a non-attending parent and children’s behaviour often differs in different relationships (such as with a mother a child may be more challenging in behaviour but less challenging with their father). Preschool children’s behaviour is often not from the same at home and at the childcare either. Bjørk’s comment regarding the use of triangulation would be more relevant with older children. Optimally, this study would have benefitted from observation

measures to verify the parent-reported outcomes. However, focusing on the valid and complete data from the participating parent is an entirely customary practice in intervention research.



Ethical aspects

The N-TIK study has undergone the standard institutional and ethical reviews typically applied in academia, including peer review for the international journal to which it has been submitted and an independent expert review committee assessment at PSI. The revised manuscript openly acknowledges the study limitations (e.g., lack of observational data) and shows that the researchers fulfilled pre-registration requirements for primary outcomes. It is untenable to suggest that this process was deliberately misleading or unethical when an independent expert committee concluded that neither the methods nor the conclusions violated good research practice.

We maintain in the strongest possible terms that Bjørk's claims of research misconduct, selective reporting and of 'overselling' findings are not substantiated. On the contrary, we believe that the N-TIK study constitutes a robust contribution to the field of parenting interventions. Parenting is the most modifiable factor that can impact children's development and functioning (Sameroff, 2010). While the reported effect sizes on children's behaviour problems may be modest, they still hold significant preventive value when applied to a broad population. Moreover, the possibility of anxiety improvement specifically in shy children is an interesting exploratory finding rather than an overly broad claim. We hope this will foster a more factual and nuanced discussion about the outcomes of emotion-focused parenting interventions and the importance of open practices in research.

References

- Belsky, J. & van Ijzendoorn, M.H. (2015). What works for whom? Genetic moderation of intervention efficacy. *Development and Psychopathology*, 27 (1), 1–6. <https://doi.org/10.1017/S0954579414001254>
- Bjørk, R.F. (2025). Når forskningsetikken nedprioriteres. *Tidsskrift for Norsk psykologforening*, 62(4), 234–238.
- Deaton, A., & Cartwright N. (2018). Understanding and misunderstanding randomized controlled trials. *Social Science & Medicine*, 210: 2–21. <https://doi.org/10.1016/j.socscimed.2017.12.005>
- Havighurst, S.S., Duncombe, M.E., Frankling, E.J., Holland, K.A., Kehoe, C.E., & Stargatt, R. (2015). An emotion-focused early intervention for children with emerging conduct problems. *Journal of Abnormal Child Psychology*, 43(4), 749–760. <https://doi.org/10.1007/s10802-014-9944-z>
- Sameroff, A. (2010). A unified theory of development: A dialectic integration of nature and nurture. *Child Development*, 81(1): 6–22. <https://doi.org/10.1111/j.1467-8624.2009.01378.x>