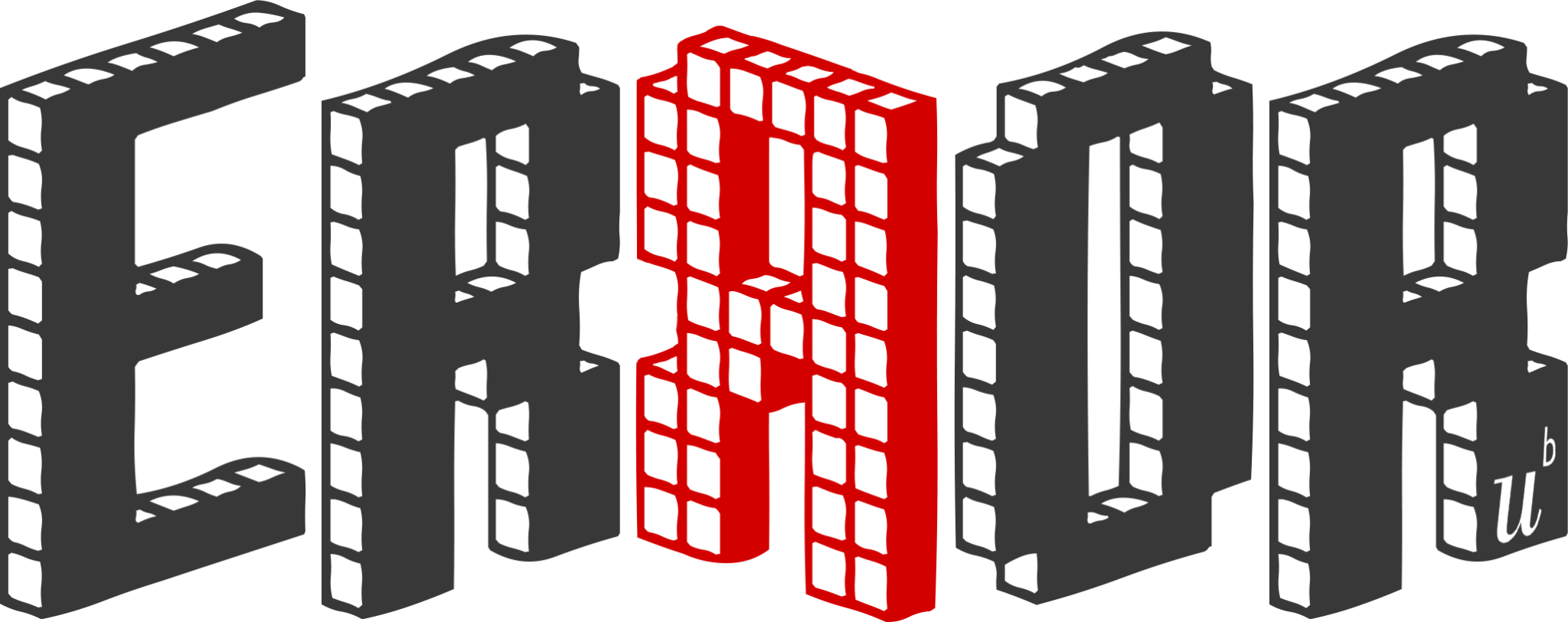
ESTIMATING THE RELIABILITY & ROBUSTNESS OF RESEARCH

ERROR RECOMMENDER REPORT

Fernbach, P. M., Light, N., Scott, S. E., Inbar, Y., & Rozin, P. (2019). Extreme opponents of genetically modified foods know the least but think they know the most. *Nature Human Behaviour, 3*(3), 251-256. https://doi.org/10.1038/s41562-018-0520-3

DECISION: Minor errors

*Recommendation by*

**Jamie Cummins**, University of Bern

5 December 2024

*Report version 1.0 (2024-05-07)*

*Recommendation template version 1.0*

*License: CC BY 4.0*

DECISION & RECOMMENDATION

Based on the reviewer’s report and the author’s response, I am returning the decision that the original article contains **Minor Errors** (see below for the ERROR rubric of error report categorisations). That is, errors that have the benefit of being detectable thanks to the presence and sharing of research materials, but whose scope and implications are minor. The detected errors do not rise to the level where I would recommend that a correction be issued. A decision of *minor errors* entails the publication of the report, authors response, and recommendation on the ERROR website ([error.reviews](https://error.reviews/)). The authors are also asked that, in future, they recognise the minor errors associated with their manuscript in future discussions of the article.

|  |  |
| --- | --- |
| **DECISION** | **RECOMMENDATION** |
| **No errors** | No additional action beyond publication of the error report and reply |
| **Minor errors** | No additional action beyond publication of the error report and reply, plus recommendation that the authors appropriately recognise these errors in future discussions of the article.  *Note that we expect most published research to fall into this category.* |
| **Indeterminable** | No determination could be made regarding the presence or absence of important potential errors. Recommendation that the authors appropriately recognise this lack of verifiability in future discussions of the article.  *Note that this is considered a less desirable outcome than “Minor errors” because in that case the errors are verifiably minor in nature.* |
| **Moderate errors** | Correction notice (minor).  *Note that numerous ‘minor’ errors might constitute a moderate error in aggregate.* |
| **Major errors** | Correction notice (major) / may warrant an expression of concern. |
| **Severe errors** | Retraction |

RECOMMENDER’S REPORT

Jamie Cummins

Firstly, I would like to thank both Prof. Fraley (reviewer) and Prof. Light (corresponding on behalf of the authorship team) for their cooperation throughout this review. Prof. Light’s response letter begins with highlighting his initial reservations towards participating in the ERROR project, including his worries relating to potential career or reputational damage. We think many authors would likely share similar worries when approached for a project such as this. We are all the more grateful for his participation and indeed we hope his positive experience with this process can serve as an example to future authors who feel a similar initial reluctance in participating.

The substance and style of the reviewer report and authors’ response embody what we hope to see more of in academic research: acceptance of the possibility that errors occur; inspection and useful discourse about potential errors that is well-documented and verifiable; and acknowledgement and suitable correction when errors are found.

ERROR recommendations are public documents whose function is to (1) communicate the presence or absence of any errors detected, (2) consider their severity, and (3) provide discussion of how similar errors elsewhere might be prevented or detected. Materials for all error reports can be found at [osf.io/fpw4r](https://osf.io/fpw4r/).

**Summary of errors detected & how they could be prevented in future**

The original article consisted of four studies investigating the relationship between extremity of opposition to certain topics (primarily genetically-modified foods, but also gene therapy and climate change), objective knowledge on these topics, and perceived knowledge around these topics. We asked Prof. Fraley to review Studies 1, 3, and 4, given that Study 2 was analysed using STATA, which Prof. Fraley was less familiar with, whereas Studies 1, 3, and 4 were analysed using R. Readers should note that Study 2 has therefore not been subjected to the ERROR review process.

As detailed in Prof. Fraley’s report, a small number of errors in the original paper were found during the review. The openly-available code associated with the manuscript was not fully computationally reproducible, with portions of code referring to objects which were not created in the script, and (more trivially) some of the R packages not loading correctly. Anyone familiar with attempting to computationally reproduce previous work will be aware that these are extremely common issues not at all idiosyncratic to the work of the original authors. In general these issues can be avoided in future work by (i) having a researcher independently attempt to computationally reproduce the results prior to the submission of the manuscript, and (ii) utilising a version management library, such as groundhog, to specify and preserve the specific package versions which were used for processing and analysis (and to document this clearly). An even more low effort solution could involve using RMarkdown files for scripts and ensuring that the file knits correctly (as undeclared objects would be caught through this process).

Additionally, the issue of computational reproducibility extended to a small number of values reported in the manuscript, where these values were erroneously reported (e.g., as Prof. Fraley noted, average age was reported in the paper as 36.6 but was actually 39.6 based on the analysis code). This small number of issues amounted to what are likely to have been typos, but can be avoided in principle by (i) having an independent researcher specifically check for such typos prior to manuscript submission, and/or (ii) writing manuscripts using a software which can directly integrate the numeric output of the code into the manuscript text, such as RMarkdown.

Prof. Fraley noted several instances in the paper where inferences about differences between relationships/conditions were made based on divergent results from two separate analyses, rather than a single analysis testing this inference directly. This is a relatively common statistical fallacy, and statistical difference cannot be inferred by the difference between a significant vs. non-significant test (or indeed, even two significant tests; Gelman & Stern, 2006) Although the inferences made by the authors in this case stood up to scrutiny with the more appropriate analyses conducted by Prof. Fraley, this can often not be the case in similar scenarios. To avoid this, authors should generally ensure that the statistical inferences they make regarding differences between relationships or conditions are mapped *directly* onto a statistical inference test, rather than based on divergences between separately conducted tests.

**Discussion of individual issues raised**

**Code functionality and computational reproducibility**

The reviewer noted issues with (i) referenced objects in the R code which were not created within the code of the authors, (ii) loading a small number of libraries, some of which were resolved by the reviewer updating his R version, and (iii) copy-paste errors between the code and manuscript. I agree with Prof. Fraley that the first two of these issues amount essentially to “hiccups” and do not substantively affect the content of the manuscript or its conclusions. The third issue affects the content of the manuscript in a small number of cases, specifically in relation to two beta coefficients which missed decimal places in Study 1, and an incorrect reporting of mean age for Study 3. Prof. Light agrees in his reply that this amounts to copy-paste errors on the part of the authorship team. Although these do represent errors in reporting, they do not substantively affect the conclusions of the paper, and therefore represent a minor error.

**Lack of direct tests for inferences of difference in relationships/conditions**

The reviewer noted that there were a number of instances where the authors made inferences relating to differences in relationships between variables based on different patterns between two analyses, rather than a direct analysis of the difference of these relationships. For instance, the authors state in their abstract that “the relationship between self-assessed and objective knowledge shifts from positive to negative at high levels of opposition” in the domain of GMO foods, but not for climate change, and subsequently compare these domains in their discussion. As Prof. Fraley stated, a more appropriate analysis would directly test the effect of “Condition” in this context to infer that something was the case in one condition and not in the other. Prof. Fraley graciously conducted this analysis (as well as others where similar issues emerged), and found that the substantive conclusions remained to be supported by the more appropriate analyses. Since the conclusions in the manuscript do not change, this may be considered a minor error in terms of its consequences for the article. It is important to note that the analytic approach is erroneous and should not be used to make such inferences. However, on balance and in the context of the conclusions not being changed, we label this a minor error in this context.

**Pseudo-causal language**

The reviewer highlighted the presence of some pseudo-causal language in the manuscript; namely, making within-person claims based on between-person analyses. I generally agree with Prof. Fraley that there is such language in the manuscript, and that this is not wholly appropriate given the analyses presented. Prof. Light similarly acknowledged and agreed that this may be a potential issue. There are certainly instances of this “pull it and it gets longer” language (as Dr. Fraley described it) throughout the manuscript, such as the final sentence “This suggests that a prerequisite to changing people’s views through education may be getting them to first appreciate the gaps in their knowledge”. In our view, even such hedged statements can mislead readers into thinking that the paper’s results mislead readers because the implicit assumptions behind such a causal inference remain undiscussed. Within ERROR, we are mindful that not all researchers agree on this point, and that statements such as these are generally common within the norms of writing in psychological research. Given all of this, I am inclined to say that this does not quite rise to the level of an error, despite the fact that readers may be misled by this type of language.

**Unresolved and unexamined aspects**

It is reasonable to expect that ERROR reviews will leave some questions unresolved. It is useful to acknowledge the potential for such issues so that ERROR recommendations do not artificially convey that they represent the final word on issues of error detection and correction for a given article. In this case, readers should once again note that Study 2 of the manuscript did not receive any review given that its analyses were conducted in STATA.

There are also dimensions of the paper which, in my opinion, could potentially be explored in more depth in the future. For instance, there are several analyses which are alluded to in the discussion section of the paper but which were not examined by Prof. Fraley. One of these analyses relates to the interaction of effects with political identification; the authors find that political identification interacted with effects for climate change but not for GM foods, but not directly test this statistical difference (i.e., echoing similar issues Prof. Fralay flagged in terms of the lack of direct tests for inferences). Additionally, the authors mention in the discussion they “re-ran the main analyses from all studies including education level as a control”, and that all key findings remained significant. However, the additional analyses Prof. Fraley ran were not tested for their robustness with such a control (and indeed, the computational reproducibility of those re-run analyses were also not inspected). To be explicit, this is not to say that any of these examples contain errors; it is simply to highlight that they were not examined in this ERROR review.

I sincerely thank both Prof. Fraley and Prof. Light again for their efforts during this process.

Jamie Cummins

Recommender for ERЯOR