Philip M. Fernbach, Nicholas Light, Sydney E. Scott, Yoel Inbar and Paul Rozin (2019) Extreme opponents of genetically modified foods know the least but think they know the most, Nature Human Behaviour, <https://doi.org/10.1038/s41562-018-0520-3>

**Error Review by R. Chris Fraley**

Here is a brief overview of my error-checking process:

1. I read the paper and made notes of elements to check. I wanted to check the pre-registration (Study 1) against what was done in the paper. I also wanted to reproduce the key results reported in the paper using the data/code provided by the authors.

2. Downloaded the data from the OSF page

3. Downloaded the R code for Studies 1, 3, and 4. (Study 2 code is for STATA, which I do not own. So, I didn't check Study 2.)

4. Ran the analysis code in base R, Version 4.3.0. I did NOT check the details of whether variables had been recoded correctly.

5. Checked reported numbers in the article against results obtained in R.

6. For any analyses I ran independently of the supplied code, I used a # Fraley comment in the newly saved script file to demarcate it clearly.

I checked all the numbers reported in the manuscript against what emerged from running the R script. I checked every number for Study 1, but focused more on the key ideas in the subsequent studies because the manuscript itself focused mostly on S1. I didn't check the bootstrapped CIs because they are bootstrapped, but in each analysis, they were in the ballpark that one would expect for bootstrapped analyses.

Total time: 5 hours.

**Simple Summary**

Almost everything checked out. A few minor errors are noted below, along with issues that are not "errors" per se, but just "concerns" that would have come up in the review process, had I been a reviewer. (But a reviewer without special theoretical knowledge in this domain.)

**(Potential) Errors:**

Extremity of opposition on self-assessed knowledge (Study 1)

\* coef of 26 reported. Should be .26

Extremity of opposition on the z-scored knowledge difference score (Study 1)

\* coef of 28 reported. Should be .28

Age Study 3

\* Average age reported at 36.6 but is 39.6 in analyses (Study 3)

**Concerns and other Minor Notes:**

Meta/Registration

\* I checked the pre-registration against the analyses and methods that were reported for Study 1. To be frank, I wasn’t expecting this to line up well because the article reports several analyses that seemed like the kinds of things that would be done to appease reviewers during the review process (e.g., analyzing the two extremity items separately and as a composite). But, in fact, virtually everything reported in the manuscript for Study 1 was part of the pre-registration too. The analyses reported that were not pre-registered (unless I missed it) were the genetics-only items from the scientific knowledge items.

\* One element that was unclear to me was the purpose of the climate change and genetically modified foods distinction. The inclusion of these two contexts, with random assignment to conditions, implies that this was an explicit factor in the original design (i.e., participants were randomly assigned to consider climate issues or genetically modified food issues). And, the way the data are labeled in the Study 1 data file (“cond” for “condition”) is consistent with this interpretation, as is the language “two-between-subjects conditions” in the pre-registration (Question 4).

However, the pre-registration doesn’t explicitly state that this factor is a factor per se in the analytic section. In practice, the authors analyze the two contexts separately rather than treating this condition as a condition in the design. The analytic section is vague on how this was supposed to be handled. Based on the registration, I was expecting Condition to be an explicit factor in the analyses. The authors make an “error” that is often made in the behavioral sciences (see Gelman & Stern, 2012, for an explanation) where they conclude that something works in one situation (e.g., genetically modified foods) but not in another (e.g., climate change) because an association emerges in one case but not in the other. But the *difference* between those associations was not formally tested. Interactions with Condition would allow for this.

\* A minor concern: The authors commonly use within-persons, repeated-measures, or causal-like language to report between-person correlations. For example, in the Abstract, the authors write that “as extremity of opposition to and concern about genetically modified foods increases, objective knowledge about science and genetics decreases…” This “pull it and it gets longer” language implies that the authors are studying the way changes in one variable for a person result in changes in another for a person. But it would be more accurate to say instead that “People who were more extreme in their opposition to genetically modified foods were less likely to have objective knowledge about science and genetics.” The latter is the “correct” language for between-person covariances. The language the authors used would be more appropriate if there were within-person covariances or experimental manipulations of knowledge or opposition.

Study 1

\* A minor concern: The intercept was not reported for the difference analyses; it was -1.28 based on the R analyses (Study 1). This may be worth reporting because the primary analysis concerns a difference between two scores and how that difference varies as a function of extremity. The coef of .28 implies that people who have more extreme scores also tend to have a larger difference between their self-assessed and objective knowledge. But, if the "starting point" is a negative value, that positive slope could indicate that more extreme people have less of a discrepancy; if the "starting point" is positive, that positive slope could indicate that more extreme people have more of a discrepancy. In other words, knowing where the cross-over point is in an interaction is helpful for interpreting regression results when the DV is a difference score.

In this case, the intercept is negative (-1.28) indicating that people with extremity values of zero tend to have a negative discrepancy (i.e., underconfidence). But the extremity variable is not standardized, so a value of zero doesn't mean anything of use on the 1 to 7 metric.

To get a better feel for what is happening, I made some plots and standardized extremity such that a score of zero would represent a person who is average in extremity.

This analysis indicated that the person average in extremity had virtually no discrepancy between objective and self-assessed knowledge. The extremity coef was positive (.50) indicating that people who were more extreme had a larger discrepancy (positive values indicating greater overconfidence).

This is compatible with the authors' interpretation. But I wanted to double-check it because the way predictor variables are centered in analyses focusing on difference scores can shift the interpretation if not considered appropriately.

\* An additional check: Because the claim is that people who have more extreme opposition to genetically modified foods show a greater discrepancy between objective and self-assessed knowledge, I wanted to perform a simple-slopes test to see if, for example, that difference is significant at the extreme end. (Depending on the form of the interaction, this may or may not pan out. But, given the additional plots I examined in the modified code, there was no reason to expect it not to.)

For Study 1, the difference is significant at both low extremes and high extremes. This latter finding is compatible with the general conclusion: That people with more extreme opposition to genetically modified foods show a larger discrepancy between their objective- and self-assessed knowledge (i.e., they report knowing more than they do). I'm not sure what the former means; the cross-over is such that the difference goes in the opposite direction. I guess, taken literally, it implies that people who do not harbor extreme opposition to GMO foods have the *opposite* discrepancy (vs. no discrepancy at all). That is, they report knowing *less* than they actually do know.

I don't think it is an "error" to leave this part out. But my intuition is that, to fully appreciate the implications of the findings, the account needs to be able to explain both ends of this continuum satisfactorily.

\* One of the core findings of this work (to oversimply it) is that extremity of opposition to genetically modified foods is positively associated with self-assessed knowledge, but negatively associated with objectively assessed scientific knowledge.

To echo the Gelman and Stern concern above: This conclusion is reached based on separate analyses (e.g., p. 252, 1st col, 1st paragraph) and noting that one coefficient is positive (and significantly different from zero) and another is negative (and significantly different from zero). An alternative way to examine this would be using mixed models where knowledge is treated as a within-persons variable (objective vs. self-assessed). I don’t think this would lead to different conclusions in this particular case; the associations are going in opposite directions (but that might not have been the case in practice).

To put together a couple of distinct concerns here: The authors have taken a mixed design where one factor is between-subjects (domain: food or climate) and one is repeated (knowledge type) and, instead of doing a mixed analysis, they have subset the data in distinct ways and analyzed those subsets separately. Not an “error” per se, but probably not ideal in other ways. But I also recognize that it is a matter of preference.

Study 2

\* I didn’t check Study 2 because I do not have a copy of STATA and the analyses for that study were done in STATA.

Study 3

\* Johnson-Neyman analyses

The package "probemod" was not compatible with my version of R, so I didn't double-check those analyses.

\* I didn't look in supplements for potential analyses. But the broad conclusion reported in the paper is that the analyses in Study 3 were consistent with those reported for Study 1. (The major difference being the order in which questions were asked.) My run of the analytic code was compatible with that conclusion. I also redid the difference analysis with standardized extremity scores and the interpretation was consistent with the authors' conclusion.

Study 4

\* Exact results are not reported in the article, but the analyses are consistent with the general conclusions reached previously. For example, people who are more extreme tend to have lower scores on scientific literacy, higher scores on self-assessed knowledge, and a larger discrepancy between objective and self-assessed knowledge.

\* For some quadradic tests, the code couldn't find s4$extremity.C, so I couldn't check those. I'm not sure if it was an error on my part.

\* No package or install code for 'stargazer.' When I manually installed it, the analyses wouldn't run. It is possible my version of R is too old.

\* I did not attempt to recreate the figures. I did create a simple version of Figure 1 just to eye-ball check things. (Looked fine.)