

PREreview: Explanation implies causation? (Myint, Leek, & Jager, 2017)

C.H.J. Hartgerink¹, Laura Kunst², and I.h.jutten²

¹Tilburg University

²Affiliation not available

April 28, 2020

Abstract

A preprint review of “Explanation implies causation?” by Myint, Leek, & Jager. Posted on bioRxiv, November 13 2017. doi: [10.1101/218784](https://doi.org/10.1101/218784)

Thank you for posting this preprint on bioRxiv! We had fun reading and discussing it. In order to structure our own reading of the paper and to not have comments disappear in the void, we compiled some of our comments for your consideration. We hope you find this feedback helpful. We also easily and happily reproduced all your tables using the provided data and code!

What is the main question the study attempts to answer?

The study presented in this manuscript investigates whether causal language in the reporting of results obfuscates the methods, goal and scope of the analysis presented. In other words: Can people distinguish between a study whose goal was inferential and one whose goal was actually causal, regardless of the language results are presented in?

What is (are) the hypothesis(es)?

The authors hypothesized a main effect of causal language on the scope of the analysis (page 2), such that inferential findings presented in causal language would be more often considered causal. Moreover, the authors hypothesized an interaction, where being presented with a “causal” answer option would show stronger scope extension than when this answer option was not available.

What techniques/analyses do the researchers adopt to test their hypothesis(es)?

The authors apply an experimental design in an online course, using a vignette about the link between smoking and lung cancer (levels: with v without causal language). Additionally, participants were either presented with causal as one of four multiple choice options, or not. As such, the design was a 2 by 2 factorial design. Students were asked to indicate the type of analysis conducted and the dependent measure is the proportion of students with the correct answer (i.e., inferential analysis)

Why is this study relevant?

Published empirical research papers oversell results and exaggerate claims by using inappropriate and out of scope statements. Rises in “inappropriate causal language” are mentioned for a few fields, such as nutritional and educational fields. Considering that humans seek order in the world, it seems plausible that this also happens in the interpretation of research findings, but that researchers might forget to keep these claims tentative. If such language affects the interpreted scope of the results, it is important as a self-reflection in order to adjust writing behavior in the future to better reflect the scope of the analysis.

Write here any general comments you might have about the research approach.

We have several suggestions for improvement that we would like to offer the authors. First we present several general and methodological points, and we then move on to discuss the rationale and relevance of the study.

Methodologically, first, the used vignette in the study is on smoking and lung cancer. We wonder whether using this example, which has a strong social/historical/cultural context of being causal, is appropriate for the point the authors are trying to make. The explanation might prime this and create stronger effects than other associations. We wonder whether using a different vignette would have produced different results (e.g., the relation between anxiety and depression or between drug abuse and impulsivity, which have less clear causal directions, or even the association between, say, the number of storks in a given country and its correlation with the number of babies born).

Second, in most scientific papers, causal interpretations of correlational studies are accompanied by ‘warnings’ that additional studies are needed to clarify the causality in a certain relationship, or that a limitation of the study is its correlational character.

Third, the vignette uses more expressive language than we anecdotally recall (“We explain” more often seems to be presented along the lines of “We expect”). This may exaggerate the problems of language.

In order to increase the external validity of the vignette, we suggest to create a different vignette (point one); create a condition in which corrective information (‘warnings’) is added versus not added (point two) and (point three) use fewer exaggerations while doing so.

Fourth, the authors do not mention how other aspects of the scholarly communication system also contribute to the overreaching in the interpretation of findings. For example, high-rejection rates at journals make editors select articles with “hot” headlines and researchers’ degrees of freedom stretch the evidence gathering process and potentially also change the scope of the analysis. Including these aspects in the (introduction of the) study would heighten the study’s relevance.

Regarding the relevance and rationale of the study, the authors describe in their introduction two ‘problems’ that may result from interpreting correlational findings as causal. They (a) list several examples of newspapers and other media outlets, which include obvious (and amusing) erroneous interpretations and (b) note that also within research, errors are made in interpreting correlational evidence as causal. The issue here is that it is unclear which of these ‘problems’ (scenario a or b) the authors focus on, and therefore what the study results implicate.

Our impression is that the authors want to focus on scenario b, thus, describing potential causal explanations for correlational findings in research as ‘wrong’, because it suggests to other researchers that the study was actually of causal nature. This is of course a challenging statement that would give rise to an interesting discussion, regarding researchers’ habits of providing potential causal explanations for their findings. There are however some issues in the manuscript that lead us to think that such a claim is not supported.

First, the introduction states that “inappropriate causal language” has been found in a few fields, such as nutritional and educational fields. However, to our anecdotal knowledge, most articles appropriately describe in their methods and discussion sections that an inferential approach has been used, and that one must be careful not to interpret the result as causal. In many research schools it is viewed as essential to provide some theoretical framework and possible (causal) interpretation of the current results, accompanied by a statement clarifying that the current data are correlational. If the authors argue that this is not happening correctly in scientific practice, more evidence for this argument should be brought forward in the introduction. Alternatively, the authors should specify to which specific scientific fields the results apply.

Second, the main audience for scientific articles consists of scientists or students with moderate knowledge on statistical methods. It seems inefficient to ask of researchers that they refrain from describing any theoretical framework that might suggest a causal relationship, because individuals without sufficient statistical and methodological knowledge might read the papers. It seems reasonable that scientific papers should mostly be

comprehensible for other scientists and students with some knowledge on methods and statistics. Given that students following an introductory class in statistics were the study population, the claim that researchers in general are interpreting correlational inferential data erroneously, seems unsupported.

This brings us to the issue of generalizability. Do you view the students as representative of researchers, the press, or the general population? The study uses a causal design by using a randomized experiment, but has a non-random sample and does not allow for inferences. One could even consider the student “sample” as being the population and deem inferences beyond this not applicable.

Nevertheless, the authors bring up an important issue in the introduction regarding erroneous interpretations of the media (named scenario b above). We suggest that your paper has more clear-cut implications for the extent to which readers who are potentially not well-versed in research and statistics (given they are students), interpret research findings as causal versus correlational, than the researcher population at large. This point of view may better match your rationale in the introduction and the examples you have included. Moreover, the students you have tested may be more representative of people-who-are-not-(yet)-well-schooled-in-statistics, such as individuals working for newspapers, than of researchers.

Finally, regardless of which focus the authors choose, we note that the paper shows there is a problem of how strongly expressive language can affect student’s interpretation of findings, but provides little framework for how to move forward. For example, it is implied throughout the paper that it is a bad thing to stretch the scope of an inferential analysis to causal (with which we agree), but does not show how we might prevent it. The manuscript seems to imply that researchers should not indicate the hypotheses underlying a certain correlation at all. If this is not the opinion of the authors, we recommend sharpening the writing to clarify what is and what is not permitted from your perspective and how to move forward.

Specific comments about experimental approaches and methods used in the study:

- It is unclear where the 68.5% comes from [page 3]. Upon reproducing, we understood where it comes from but it distracted us during the reading.
- During the reading it was really difficult to understand how 20,256 students resulted in approximately 11,000 included in the tables. If this is due to the exclusion of re-takers of the test, then please indicate the difference in the text.

Specific comment/notes about figures in the paper :

We really enjoy Table 1, and the CC BY helps for reuse in classes. What we did during the reading of the paper was draw out the percentage correct in a simple plot, to visualize how the conditions relate to each other. This also facilitates the interpretation in light of the hypotheses, which seem to imply a main effect and an interaction. We think it would be helpful if you would add such a visualization, because in our experience the interpretation is easier to discern visually.

Any additional comments:

- “We explain, we think” [page 3]
- “The size of the effect is must smaller” incorrect? [page 4]
- ‘To confirm our results.. ’ is it really a confirmation, and is this the reason its done? [page 5]
- ‘the data is visualized ... ’ -> data are visualized [page 5]
- GitHub is not a sustainable hosting service — you could archive the contents for preservation at Zenodo via the GitHub release link.