PCI Registered Reports #188: How Intelligence Interviewees Mentally Identify Relevant Information

Dear Zoltan,

Many thanks for the opportunity to revise and resubmit the captioned registered report. We appreciate the reviewers' continued efforts to help us improve. This revision addresses the latest concerns. Next follows our point-by-point response to the editorial comments. We have included the comments verbatim (in black font), and our response follows in red font.

Sincerely, Authors

Table of Contents

Editor:	1
Anonymous reviewer, 15 Aug 2022 20:26	2

Editor:

1) In your Design Table, list each hypothesis as a row, so it is clear exactly which test addresses which hypothesis and what follows from possible outcomes in each case.

In truth, it is challenging to execute this request and produce a tidy, reader-friendly, table. This constraint is due to the fact that the hypotheses are interrelated. Supporting or rejecting any of the predictions will assist in interpreting the other predictions and refining the theory. Furthermore, there are no short intuitive names that explain, at a glance, the models slated to test the respective hypotheses. So, we explain the tests broadly and refer the reader to the analysis plan and other tables which contain the precise specifications presented more tidily.

That said, we sought to find a happy medium with the revised table (p. 32 without track changes). The revised table lists the core hypotheses and we summarize the implications of each hypothesis for the theory at large.

2) Be clear about when you will not conclude anything, i.e. when the ROPE lies entirely within the HDI (which may happen if your maximum N is reached before the HDI acquires a smaller width than the ROPE).

On page 22, we have now noted explicitly that we will refrain from drawing any conclusions when the ROPE lies entirely within parameters' HDI.

3) You refer to the bottom limit of the 95% CI for particular past studies in terms of minimal plausible effects; what if you looked at the bottom limit of the CI for the combined effect from the studies you list? Just a thought. You could go with a 90% CI. Otherwise, it is not clear, as you also indicate, you have really found a minimally interesting plausible effect size. Given that an interesting effect may be smaller than the limits of your ROPE, due caution should then be reflected in the conclusions that follow (e.g. explicitly indicating, including in the abstract, that effects smaller than a certain amount are possible).

Suppose we decided to use the bottom limit of the 95% CI for the combined disposition effect from Lorson, Cummins, et al. (2021) and Lorson, Rohde, et al. (2021) instead of examining the smallest possible effect (i.e., 0.54). Based on a combined 95% CI of [0.24, 0.83], we would yield a nearly identical ROPE as the one we specified ([-0.24, 0.24] instead of [-0.25, 0.25]). That observation leads us to believe our specified ROPE [-0.25, 0.25] is prudent at this time. That notwithstanding, we acknowledge that we run the risk of missing plausible effects that are smaller than the limits of our ROPE. We have acknowledged that potential limitation on page 27 and will interpret the findings accordingly.

Anonymous reviewer, 15 Aug 2022 20:26

I thank the authors for their thoughtful responses and edits. Below are additional/remaining comments I have after reviewing the revised manuscript.

1. I appreciate the authors' housekeeping note in the first paragraph. However, I believe that it can be moved to a footnote so that it does not detract from presenting the core aim of the paper.

Thanks for this suggestion, we have organized all housekeeping notes under footnote 1.

2. Concerning the section Situating the Present Theory: I suggest that the authors consider reorganizing the section. They bury the most important part of the section— hearing what is being questioned from a question—after a paragraph about the CI. I suggest that the authors start out with this important topic, and just give a brief nod to the CI within this section.

3. It might be worthwhile for the authors to briefly mention some of the research on question types (e.g., Griffiths et al., 2011; Milne & Bull, 1999) in the Situating the Present Theory section. This topic is also something that the authors should consider as they draft their Discussion section in the future.

This response tackles points-2 and -3. We agree and have culled the part where we describe the CI's details. The discussion now transitions seamlessly to the crucial aspects: hearing what is being questioned from a question. To accommodate point-3 without reintroducing the limitations point-2 brings up, we mention that strategic questioning is also a macro level issue and cite Griffiths et al., 2011. That article deals with interviewers' strategic selection of question-types to elicit information.

4. Regarding your use of Prolific: Are you using a balanced sample? A representative sample? Neither?

Prolific's panel is international. Thus, we will use convenience sampling, Prolific's primary sampling strategy. Study places will be filled on a on a first-come, first-serve basis. But we will balance sampling across sex. We have now included that we will balances across sex (p .16)

5. On page 22, the authors state that there are 3 outcome levels in Study 1; however, is there not 4 levels (i.e., the authors did not list "I cannot determine what the interviewer wants to know")?

Many thanks for drawing our attention to this oversight. Yes, in truth, there are 4 outcome levels in Study 1. But for our simulations, we drew on the outcomes that were directly relevant for the theory. We included the response: "I cannot determine what the interviewer wants to know" to mirror the "no comment" option in Study 2. That measure will also provide insight into our design's robustness particularly the efficacy of our high-vs. low-worthwhileness manipulation. All things being equal, high-worthwhileness questions should better reveal what the interviewer wants to know. The "no comment" option is directly relevant to our theorizing in the utterance experiment (Study 2), given that interviewees can choose to remain silent whether or not they can determine what the interviewer wants to know.

We meant to include the adjective "relevant" to highlight our clarifications above and have now done so (p. 22). We chose that presentation style for conciseness. We will raise these issues in the discussion.

6. Something for the authors to consider when thinking about the Model 2 predicted results (pg. 25): Resistant subjects may opt for bare utterances over no comment to appear cooperative, and especially so when asked high-worthwhileness questions.

Absolutely! We are inclined to agree with this observation and this is a matter we might have to revisit in the discussion and future studies. For now, we want to specify the hypothesis to be strictly in line with the current theorizing. The results will assist us in specifying more precise hypothesis in the future.

7. I thank the authors for addressing the compound questions and vagueness of some of the scenarios! My remaining comment is for question 2 – the authors might consider how to frame the low-worthwhileness question in a manner that focuses in on time.

Agreed, we have now revised question 2 to focus on time: "Do you have any information about *when* the gang deals drugs?"

8. I very much like the authors incorporation of Scharff into their future Discussion section! However, I still feel that the authors should consider including a little more (not

necessarily a whole section) on the literature that backs their hypotheses for Study 2 (differences in utterances between cooperative and resistant subjects). Their hypotheses make sense to me (as I am familiar with the literature), but may appear to have little backing to readers who are not entrenched in the investigative interviewing literature.

Thanks for drawing our attention to the potential for readers to flag the reasoning behind Study 2's hypothesis. We have now added supporting literature on page 16 (paragraphs 1 and 2).