This Stage 1 submission sets out a plan for a study to examine whether feedback about commission errors (remembering something that did not happen) and omission errors (forgetting something that actually happened) will influence participant’s criterion placement, and also the correspondence between their beliefs about their performance and their recollection. I write this review from the perspective of someone who is familiar with the methods used in this paper, including signal detection, and the analyses proposed.

Overall, the logic and methods seem appropriate to the question. The scientific validity of the question is justified, the logic of the hypotheses tracks, and the analysis is appropriate. I do, however, have some comments on the clarity, and on the exclusion criteria and outcome-neutral conditions, which I set out below:

Clarity – This is a complex area, and consequently the introduction needs to give more explanation of were some key ideas required for a reader to understand the study. I believe this could be addressed by focusing on two parts of the introduction. First, the 2nd paragraph on p.4 introduces some key ideas which could be fleshed out and given some additional examples, to help readers to follow the central ideas of the paper. Second, on p6 the idea of a decision criterion is introduced, but a high degree of knowledge about SDT is assumed here. I don’t think that SDT needs to be explained in full, but at least a brief explanation of what the criterion is doing, and a few more examples of why it might move in response to different circumstances would help the reader to follow the logic of the experiments. Given that the experiment is making some fairly fine distinctions between different types of rating that sound similar (but aren’t) this detail is important so that readers can fully understand the difference between the ratings, and how they each relate to the criterion.

Minor Clarity points:

1. The words “filler” and “new” seem to be used interchangeably in the method, with no description of how they relate to each other. Provided that the terms “old” and “new” are well described, I see no reason why the word “filler” needs to be used at all, which will help with readers following the method.
2. How long did participants have to complete the second part of the experiment? What’s the potential range of the gap between the first and second part?

Exclusion criteria – This strikes me as an experiment which requires participants to be both naïve as to the purpose of the experiment, and also at least to some degree uncertain as to whether their beliefs are accurate. Thus, in the exclusion criteria, it would be appropriate to consider ceiling and floor effects, as well as participant’s responses to the open question about what they thought the experiment was about.

In summary, I think this experiment should be conducted, but the bar lowered for how much existing knowledge is required to understand the design. I look forward to seeing the results.

Signed,

Dr Greg J Neil