ORIOLE MANUSCRIPT REVIEW HISTORY
REVIEWS (ROUND 2)

Editor Decision Letter

Thank you for submitting your manuscript to JCR. The same review team who read your previous submission read this version as well. I have carefully read your manuscript and reviewer responses, and digested the reviewers’ comments. As before, I appreciate the clarity of the reviewers’ responses. The AE has provided clear and specific advice for moving forward. This letter provides my decision regarding the paper and the reasoning behind it, taking into account the clear guidance of the review team.

We all appreciate how thorough and responsive you have been to the previous comments. The paper is much clearer, the contribution is more solid, and the results are convincing. Congratulations on an excellent revision! At this point, I would like to offer a conditional acceptance of the paper. Just a few issues remain. Since they are quite clearly articulated by the AE (see the next to the last paragraph of his/her report), specific and doable, I will not burden the reviewers with another re-read. Instead, I will ask only the AE to take another look at the paper. I would like you to provide a set of comments to the AE that indicates how you have addressed the comments that remain.

Congratulations on a fine revision. You have come a long way in developing it and I hope that you can return that final version in relatively short order.

AE Report

Comments to the Authors:

All three reviewers, the trainee reviewer, and I agree that the authors did an excellent job of revising this manuscript. Reviewer A recommends that the proportion recall measure be used in all experiments. Reviewers B and C recommend explaining how attempt-to-align false recall leads to poor choices. Reviewer C also recommends explaining how relief from accountability operates differently for experts vs. novices. Reviewer C also recommends explaining the responses of novices in greater detail. Attempt-to-align false recall is one type of memory intrusion (see Wyer and Srull 1989, Memory and Cognition in its Social Context, Hilldale, NJ: Erlbaum), and intrusions are interesting because they involve schematic information processes and because they demonstrate the high degree of willingness of experts to go beyond the information given. These reconstructive inferences are frequently confused with perceptions (Johnson, M. K. 2006, American Psychologist; for marketing implications of source monitoring,
see Kardes 1988, Psychology & Marketing). The focus on alignability is also consistent with extensive research showing that alignable attributes are weighed more heavily in similarity and choice, relative to nonalignable attributes (Markman and Loewenstein 2010, Journal of Consumer Psychology).

Following the reviewers’ advice, I recommend using the proportion recall measure in all experiments, and explaining how attempt-to-align false recall leads to bias and error in decision making. Attempt-to-align false recall could bias choice by decreasing the influence of unique features (Houston et al. 1991, Journal of Experimental Social Psychology), and by increasing the spreading of alternatives (Liberman and Forster 2006, Journal of Experimental Social Psychology). I also recommend relabeling “source confusion” simply “confusion,” consistent with Burke and Srull (1988, JCR) and to avoid confusion with the term “source monitoring” (Johnson 2006, American Psychologist). Finally, in the general discussion, it would be useful to consider the implications of the results with respect to the SCAPE model (Whittlesea et al. 2008, in Handbook of Consumer Psychology, eds. Haugtvedt et al., New York: Psychology Press), which suggests that all memory experiences involve the operation of reconstructive inference processes. Loftus and her students also make this argument (Braun 1999, JCR; Braun et al. 2002, Psychology & Marketing), but see Wyer (2008, in Handbook of Consumer Psychology, eds. Haugtvedt et al., New York: Psychology Press), for an alternative perspective.

Again, the authors did an excellent job of revising this manuscript, and the contribution of this manuscript can be increased even further with a few relatively straightforward revisions.

Reviewer A

Comments to the Authors:

I was positively disposed to this manuscript after the first submission and my evaluation has improved further. My primary concern with the original submission regarded the rationale and clarity of the coding scheme of “spontaneous false recalls” and an “attempt-to-align false recalls,” a concern that was echoed by other individuals on the review team. In response, the authors developed a much more parsimonious coding scheme focused on “incorrect value recalls,” “source confusions” and “attempt-to-align false recalls” and replicated their previous findings with this new scheme. The authors also addressed all of my other concerns sufficiently.

One concern that was voiced in the reviews (by Reviewer C) was that the observed effects might not persist if the authors assessed alignment false recalls as a percentage of overall recall as opposed to assessing them with a simple count. Although I did not catch this issue on the first submission, I strongly agree with Reviewer C and the AE that this is a significant concern with the previous submission. Given that, I was pleased that the entirety of the results (including the mediation tests) replicated with this measure of proneness to alignment false recalls. For the sake of space, the authors only reported these percentage results for experiment 1 which I certainly understand given the overall length of the manuscript. However, I was left wondering whether the percentage results might be the more appropriate focal dependent variable throughout the manuscript.
Well done.

**Reviewer B**

**Comments to the Authors:**

I was very impressed and pleased by the detailed and responsive reviewer notes. It is clear that a tremendous amount of effort was undertaken in order to appease our concerns. As a result, I believe this revision is a significant improvement over the prior draft. In fact, I have only one remaining concern. This concern has to do with whether the expert’s false memories are in fact more likely to lead to negative versus positive outcomes. The paper assumes that these memories lead to negative outcomes (and indeed this is shown in study 3 as a decrease in evaluation of the “superior” option). However, the conclusion of study 3, that experts’ false memories lead to a decrease in their evaluation of the superior option, assumes that the information provided for Option B is the only information that could correctly be used to determine that in fact option B is objectively superior to option A. However, experts are much more likely to make inferences regarding what could be considered missing information (layman do not have the requisite knowledge to form these inferences), and these inferences may be correct. Hence, experts’ false memories may be partially composed of these correct inferences that were made during encoding. If so, then perhaps experts were in fact correct to decrease their evaluation of Option B relative to non-experts, and it is the inferences in addition to the presented information that dictates the extent of Option B’s superiority over Option A.

For example, in the stimuli used for Study 3 (as shown in Appendix 2), I assume recalling some level of video resolution for Option B would be considered a false memory due to an attempt to align features since 1080p video resolution was mentioned for Option A, but nothing about video resolution was mentioned about Option B. However, based on their knowledge, an expert may infer that because Option B did not have HDMI Output available, it likely had something less than 1080p video resolution because 1080p resolution and HDMI output availability are two attributes that usually go hand in hand (note – I know little about this technology stuff so this specific example may not be correct, but hopefully you get the main gist of what I am getting at). In other words, if HDMI output availability and 1080p resolution co-vary, then wouldn’t it make sense that experts infer one attribute based on another? Hence, it is quite possible that inferring option B has something less than 1080P resolution would be a correct assumption since it did not have HDMI output availability. It is also possible that this inference was falsely recalled as being presented. This in turn would make experts down-grade Option B’s superiority over Option A, which is what you found in Study 3. Further, making this assumption would: 1) not necessarily rely on an attempt to align features across options but rather simply knowledge of covariation across attributes, and 2) may in fact be a reasonable assumption. Could you provide any insight, conceptually or analytically, to address this concern?

**Reviewer C**

**Comments to the Authors:**

Overview
There is a lot to like about this revision: the positioning of the paper is clearer (especially as regards accountability and expertise); the recoding scheme and elimination of terms (e.g. spontaneous false recalls, accuracy) makes the theoretical development clearer; the elimination of Study 1 improves the contribution/length ratio.

The author’s revised treatment of the ‘experts remember more’ idea is ok with me (i.e. it is dealt with in the first study but not afterwards). I continue to feel that an objective measure of accountability would have been superior, but I don’t see it as a mission-critical issue.

There are still a few outstanding concerns:

As-Yet-Unresolved Issues

1) Clarify accountability yet again. Although it’s not hard to accept that experts might naturally feel more accountability, it is less clear at what form of accountability – x-expertise interaction we might expect. (If the base/control level of accountability is high, we might expect novices to react more strongly to an accountability manipulation; and, in the absence of any actual knowledge of the category, this reaction might be a negative tone.) The answer presented on p. 10 makes sense to me (a low-accountability manipulation is to RELIEVE experts of accountability), but I’d still like to know how “relief from accountability” works differently for novices.

2) Related to this, H3b, H4a, and H4b, have nothing to say about how you expect novices to react. Can you clarify this? It would be in keeping the paper’s positioning.

3) H4a and H4b are still worded rather ambiguously: “By encouraging a focus on the decision outcome, expert’s attempt-to-align false recalls will be reduced.” Do you mean, “Reduced vs. when the focus is on process” (implying you are going to show different effects for process accountability v. outcome accountability)? Or, do you perhaps mean “the difference between experts and novices will be reduced”? It’s unclear (particularly because you have not had anything to say on how novices will react to the manipulations implied in H4). In H4b, what do you mean by “experts’ product evaluations will improve”? Do you mean the average evaluation will be more positive? Do you mean that they will be more ‘accurate’? Again, unclear. Please help!

4) I still feel that the paper’s ‘complaint’ about experts, while technically valid, may be limited in its implications - in the ‘real world,’ much of what experts incorrectly recall to have been present in an ad might be perfectly true. The hotel ad did not say, “we have electricity 24 hours a day,” but it seems a bit petulant to tell an expert traveler he/she is wrong to have assumed that electricity will be present. Clearly, the traveler that actually remembers that ad as having made a claim about electricity is incorrect – but pragmatically speaking the expert is better informed and has a more accurate idea what to expect than the novice – because his/her schemas are probably right most of the time. When, then, if ever, is an experts’ tendency to resort to remembered schemas likely to get him/her in practical trouble? [Presumably in situations where a product category is rapidly changing, or perhaps when an
expert’s schema is geographically inappropriate – an expert (American) traveler on safari will be more surprised to find he often cannot charge his laptop in mid-afternoon even in high-end resorts than will a novice.] I think that some discussion of the likelihood and conditions under which this effect will “matter” in daily life would be a useful addition to the closing discussion section.

Overall, a very careful revision, thanks.

Trainee Reviewer

Comments to the Authors:

This manuscript was very enjoyable to review. The authors did a very thorough job addressing both my concerns and the other reviewer’s concerns in the latest revision of the paper. I have no major concerns about this paper and have only a few minor concerns that I think should be addressed before this paper moves to publication.

Minor Concerns

- Both the first and second example in the paper related to videogame consoles, specifically, a comparison of the features from the Xbox 360 and the Playstation 3. While these are strong examples, I feel that they may take away from the reader applicability of the false memory product comparison construct to other contexts. Are there any other product contexts that you could use? Specifically, you could examine phones, music players, or a car. In addition it may be strong to look at a product category that is non-technical such as buying a mattress. I feel it is particularly important to vary BOTH of these examples since for E1, E2, E3, and E4 you use the exact same context. I am not going to ask that you go back and change your experiments so you use different stimuli, but on the front end can d the conclusion you can easily change out the examples.

- I know that you manipulate accountability but for a future study can you manipulate expertise? Can you teach a novice to become an expert by helping them create a more developed schema? Does a new “expert” behave the same as an old “expert”? I bring this up because of the Patel and Groen, 1991 study with the 3rd year residents and doctors. Many would suggest that a 3rd year resident is a “new expert”, they certainly know much more than a novice, but they know less than an “old expert”. In what contexts would a new expert behave like a novice and what contexts would they behave like an expert? Or is there a completely different behavior that they would engage in?

- Do you have any citations to support your contention that for memory-based product comparison tasks you expect accuracy to diminish? Has anyone else examined this?

- Is a focus on the process only helpful to novices – or only in the product comparison context is it only helpful to novices? Similarly, is a focus on the end result only negative to novices? Are there any other contexts that we expect these effects to occur? The reason I am focusing
on this is that this may be one of the most important parts of this paper. I would try and highlight this a bit more in the introduction.

The experiments were very well done. The concerns that I want addressed were all from theory development and the introduction. Good luck in your future research!