June X, 2023

Bernard Hommel, PhD

Action Editor

*Psychological Research*

Dear Dr. Hommel,

My co-authors and I have submitted a revised version of our manuscript PRPF-D-21-00324 “Item-Specific and Relational Encoding are Effective at Reducing the Illusion of Competence.” We are encouraged that our reviewer found this version of the manuscript to be “greatly strengthened” relative to our previous experiment while positively commenting on the quality of the writing and the scientific rigor of our experiments. Below, we have provided a set of point-by-point responses to each comment. In doing so, we have focused on revision on clarifying the motivation for this study while also streamlining the Introduction and Results section. To facilitate review, all major modifications to the manuscript have been made using blue font. We also include page numbers when referencing specific changes to the manuscript. We appreciate you inviting a revision and hope our revised manuscript is now suitable for publication in *Psychological Research*.

Sincerely,

Mark J. Huff, PhD

School of Psychology

The University of Southern Mississippi

Email: mark.huff@usm.edu

Phone: (601) 266-5411

Fax: (601) 266-5580

Cc:

Nicholas P. Maxwell, PhD

Midwestern State University

Emily Cates

The University of Southern Mississippi

**Action Editor:**

Dear Dr Huff,  
  
Thank you for submitting your manuscript, "Item-Specific and Relational Encoding are Effective at Reducing the Illusion of Competence", to our journal. I had hoped to secure two reviews from experts of your research area. So far, I have only received the one review that you find below, while I'm waiting for a second for way too long and a third reviewer just let me know that he or she cannot make it for personal reasons. Hence, I decided to base this action letter on the one review I have and on my own reading.  
  
As you will see, the referee is positively inclined towards the study and the manuscript but has quite a number of substantial objections. Objections that I share. In particular, the theoretical motivation for the study is underdeveloped, and the degree of novelty isn't clear from the overview. I think that the first point connects to another issue the reviewer raises, namely, the lack of structure in presenting the findings. If the theoretical frame would be tighter, the reader would have a better idea about which findings are more, which findings are less important. The referee and I feel that these issues can be addressed, the problems can be fixed. Therefore, I would like to invite a revision that convincingly addresses the remaining issues, together with a cover letter that explains how this was achieved. My plan would be to send the revision to the present reviewer, if possible.

***Response:*** We have revised our manuscript in accordance with both your feedback and our reviewer’s comments. In doing so, we have focused our revision on streamlining the Introduction, clarifying the rationale for this study, and simplifying our Results section. Regarding the Introduction, we have shortened our discussion of previous studies investigating JOL accuracy (pg. xx), reworked our discussion of LOP effects on pg. xx to explain our rationale more clearly for using deep encoding tasks to reduce the illusion of competence, and have remove our discussion of calibration and resolution (pg. XX of our initial submission; see our responses to Comments 1 and 2). This section has subsequently been replaced with a new section which summarizes and explains the motivation for the present study (see pg. xx of our revision). In doing so, we believe that our revised Introduction more clearly communicates the rationale behind our study while highlighting the novel contributions of this research.

Regarding our Results section, we have removed both the resolution analyses and the cross-experimental analyses. In doing so, this allowed us to focus our results primarily on changes to the illusion of competence as a function of encoding group. However, for completeness, we now report these analyses in the Appendix (pg. xx), given that some readers may be interested in these findings (Comment 5).

**Reviewer 1:**

The present studies examined the illusion of competence, when it occurs, and how to reduce it. The researchers tested how item-specific and relational processing tasks affected cued recall performance, JOLs, metacognitive calibration, and metacognitive resolution compared to read-only studying. The learning materials were 4 different types of word pairs: unrelated, strong forward association, strong backward association, and symmetric. The results were largely consistent across both experiments. Among backward and symmetrical pairs, the illusion of competence emerged in read-only study condition and was reduced or eliminated when participants studied with relational or item-specific processing. The improvements in calibration can largely be attributed to improvements in recall, not changes in JOLs. And although calibration improved with relational and item-specific processing, resolution decreased.

This paper is a revision of an earlier submission. This new version is greatly strengthened by the addition of Ex. 2, which replicated Ex. 1 and found a highly similar pattern of results. Ex. 2 also used a think-aloud protocol during the study phase to ensure that the item-specific and relational processing instructions successfully affected how participants encoded the pairs. I also commend the authors for providing such detailed analyses (JOLs, recall, calibration, resolution, and cross-experimental). Despite these improvements to the manuscript, I have several concerns that temper my enthusiasm for the manuscript and its publication in its current form. I believe these issues could be addressed with writing, though, and would not require additional data because the Ex. 1 and 2 were carefully designed. I hope the authors find some of these comments helpful and I wish them the best of luck in pursuing this work.

**Comments – Theoretical Motivation**

One of my biggest critiques is that I struggled to understand the theoretical motivation for the studies. As a result, it was difficult to conclude how the results meaningfully advanced the field’s understanding of associative learning and metacognitive monitoring. Below are some questions/suggestions relevant to this point.

**Comment 1:** Theoretical Motivation for the Study: The stated goal of the studies was to test “whether encoding tasks that emphasize the shared or distinctive characteristics of the word pairs through relational and item-specific encoding tasks, respectively, can improve the predictive accuracy of JOLs” (p. 4). Similarly, the study was interested in “investigating how item-specific/relational encoding processes influence the relationship between JOLs and recall” (p. 10). I was left wondering, though, why we want to investigate this. I think part of the problem is that the research is framed in terms of researching cue-target relations. Indeed, the first sub-header in the Intro is “The Effect of Cue-Target Relations on JOL Accuracy.” As a result, the Intro left the impression that the goal was to see what happens when you combine popular paradigms (e.g., LOP) and materials (e.g., forward vs. backward pairs), rather than how these paradigms and materials can be used to investigate specific questions about memory and metamemory processes. This is not to say there isn’t an important research question being addressed here. The Intro just needs to be adjusted to better highlight the deeper theoretical questions/hypotheses being tested about the Illusion of Competence?

***Response:*** Thank you for taking the time to review our manuscript. Our previous submission included a lengthy discussion of cue-target relations in the Introduction, given that associative direction (i.e., forward vs. backward associations) is a driving factor behind the illusion of competence.

We have streamlined this section of our Introduction to focus less on cue-target relations while shifting the focus to metacognitive illusions and, specifically, the illusion of competence. In doing so, we have updated the first section heading (pg. 4) to read “The Illusion of Competence.” We have also reframed the first two paragraphs of this section (pgs. 4 and 5) to introduce the illusion of competence, with the discussion of cue-target relations limited to describing how different types of relations affect the correspondence between JOLs and recall, thus producing the illusion of competence for backward paired-associates.

**Comment 2:** Organization of the Intro: For the most part, the Intro does a nice job of summarizing prior research. Even though the writing was relatively clear, I struggled keeping track of how the present experiments were similar to and different from this prior research. I was having a hard time figuring out what the reported experiments did that prior studies have done, and what was new. I think the issue was mostly that there were just so many topics covered in the Introduction: predictions vs. post-dictions, calibration vs. resolution, levels of processing, item-specific vs. relational encoding, generation, timing of JOLs, JOL reactivity, experimenter vs. self-paced studying, cued recall vs. recognition, calibration plots, and more. I’m wondering if the Intro could be reorganized and streamlined to more clearly address what prior research has done, what was missing from that research/what open questions remain, and how the reported experiments extend prior research to address these gaps in the literature.

***Response:*** We appreciate your feedback. To be comprehensive, our Introduction included a detailed overview of previous research assessing the accuracy of JOLs. Starting on pg. 4, we have streamlined our discussion of previous studies investigating JOL accuracy. Specifically, we have removed the discussion of post-dictions from our discussion of Arbuckle and Cuddy (1969). We have also reworked this paragraph so that it is more closely linked with subsequent studies on JOL accuracy.

Next, we have condensed our discussion of previous studies investigating the illusion of competence. Specifically, we have combined the paragraphs discussing results from Castel et al. (2007) and Maxwell and Huff (2021), which now provides a more straightforward discussion of previous studies replicating and extending Koriat & Bjork’s (2005) original illusion of competence study. Additionally, we have expanded upon our rationale for using the item-specific/relational encoding tasks (pg. 7).

Additionally, have clarified previous research combining levels-of-processing with JOLs. Specifically, we have streamlined our discussion of Tekin and Roediger’s (2020) work while expanding upon the motivation for the present study. Specifically, we now note on pg. x that few studies have investigated the effects of encoding manipulations on JOL accuracy, and to date, no study has assessed the influence of item-specific and relational encoding tasks.

Finally, regarding our discussion of calibration and resolution, we have removed this section. Instead, we have replaced this section with a new heading (The Present Study, pg. x) which briefly summarizes our motivations for this study while also introducing calibration plots as a tool to assess the illusion of competence.

**Comment 3:** Calibration vs. Resolution in the Intro: Related to points 1 and 2 above, I got a bit lost in the discussion on factors affecting calibration vs. resolution. I didn’t understand the distinction between factors that influence calibration and resolution: “Whereas calibration is strongly influenced by factors that affect the magnitude of JOLs/recall, resolution is primarily impacted by factors influencing retrieval, including testing, practice, and timing” (p. 12). What is the difference between factors that affect the magnitude of recall and factors that affect retrieval? After all, testing, practice, and study time affect recall. It also seems a bit circular to say that factors that affect JOLs/recall affect calibration since calibration is just JOL – recall. More importantly, I don’t know if this discussion of calibration vs. resolution helps motivate the present research. My understanding is that the illusion of competence research is all about calibration. Why was resolution also reported (besides the fact that it could be calculated with the data collected)? Was there a specific research question that could only be answered with resolution? Was there a specific hypothesis being tested by comparing the effects of the experimental manipulations on calibration and resolution? Perhaps the calibration vs. resolution discussion could be shortened/removed from the Intro, especially since the list of factors influencing resolution aren’t manipulated in the reported experiments.

***Response:*** We initially included a discussion of calibration and resolution, as it was unclear whether both types of JOL accuracy would be affected by item-specific and relational encoding. However, in an effort to improve the focus of the introduction and simplify the results section, we removed this section (please see our response to Comment 2).

Given our removal of this section from the Introduction, all analyses assessing resolution are now reported in the Appendix (pgs. xx – xx; please see our response to Comment 5).

**Comment 4:** Recall vs. JOLs in the General Discussion: What are we to make of the fact that the reduction in the illusion of competence was driven by an increase in recall, not changes in metacognitive judgments? The results of the present study essentially amount to a demonstration of a LOP effect, and its moderators. To me, this would represent a relatively modest contribution to the literature. Therefore, I encourage the authors to address what these results tell us about how people monitor their learning, the cues they use, how they combine these cues, etc.?

***Response:*** Our finding that item-specific and relational encoding strategies largely reduced the illusion of competence via improved recall is consistent with previous research suggesting that deep encoding tasks modify memory without modifying JOLs. For example, as noted on pg. XX, Tekin & Roediger found that deep encoding tasks improved recognition but did not modify JOLs. Given that the present study used cued-recall, our findings show that this pattern is not limited to recognition.

However, from a cue-utilization standpoint, it is a bit surprising that providing item-specific and relational encoding strategies largely did not affect JOLs, particularly given that relational encoding would be expected to highlight pre-existing cue-target associations that participants use when forming their JOLs. However, we note that across experiments, both encoding strategies increased JOLs for unrelated pairs, while JOLs on related pairs were largely unaffected. Thus, it may be the case that for related pairs, the pre-existing cue-target associations are already decidedly strong, and benefit little from the requirement to engage in additional encoding. However, given that intrinsic cues are weaker in unrelated pairs, both encoding tasks may be more likely to strengthen these cues. We now discuss this point on pg. xx in the General Discussion.

**Comments – Results:**

Another issue is that I had a really difficult time following the results. It’s not that they were poorly described; the writing was clear. It’s just that there were so many analyses it was hard to keep track of them all and make a mental model of which were most important for the research question as I read.

**Comment 5:** Could the authors create some sort of summary table for each experiment that verbally captures the key comparisons and results? It doesn’t have to be a table, but I think the manuscript would greatly benefit from a quick way to get the most important results.

***Response:*** We agree that the present study contains an abundance of analyses. For each experiment, we initially included both a traditional analysis of mean JOLs and recall, calibration plots and their subsequent analyses, and an analysis of gammas to assess changes in resolution.

In our revised manuscript, we have sought to simplify the results section by moving the gamma analysis in each experiment to the Appendix (see our response to Comment 3). We have also moved the cross-experimental analyses into the Appendix (pg. xx; see our response to Comment 6). Thus, each results section now simply contains our analyses of means and calibration plots.

**Comment 6:** It might help to put cross-experimental in a supplement or appendix. The cross-experimental analyses aren’t directly relevant to primary research question, but rather function to show that the inclusion of the think-aloud protocol didn’t dramatically change the overall pattern of results. I’m glad they were conducted but perhaps the authors could note the this in the manuscript and refer the readers to the cross-experimental analyses elsewhere.

***Response:*** Thank you for this suggestion. As you noted, we included the cross-experimental analyses for completeness, as we were initially concerned that our inclusion of the think-aloud procedure in Experiment 2 might alter participants were encoding the cue-target pairs and affect JOLs and/or recall relative to Experiment 1. To streamline the results section, we have followed your suggestion and moved the cross-experimental analyses to the Appendix. The full results for this set of analyses are now available on pg. xx.

**Minor Comments**

All comments, including spelling and grammatical corrections, have been addressed. We appreciate your attention to detail.