Gmail

Ian Hussey <ianhusseysipod@gmail.com>

Follow-up based on review of PSP-A-2019-0728

Payne, Keith <bkpayne@email.unc.edu> 25 November 2019 at 16:02
To: Jamie Cummins <Jamie.Cummins@ugent.be>
Cc: Sean Joseph Hughes <sean.hughes@ugent.be>, ian hussey <ianhusseysipod@gmail.com>

Dear Jamie, Sean, and Ian,

I think the design you propose would be an interesting experiment to do, but I don't see how it would provide evidence about missattributions or confabulations. It would provide evidence about the accuracy of predictions. If you find that subjects are above chance at predicting when they will be influenced (and hence that subjects who predict they will be influenced show larger priming effects) it suggests that people can make such inferences based solely on their affective response to the primes. They can't report on any (mis)attributions about the target because it has not yet been resented. So more intense affective reactions are a basis for inferring that their target evaluations will be influenced. I would think this provides an additional mechanism, beyond your awareness account, by which people might reach judgments above chance. Wouldn't that be an alternative account, as opposed to more evidence for your direct awareness account?

Best,

Keith

Keith Payne, Ph.D.

Professor and Director of Graduate Studies

Department of Psychology and Neuroscience

From: Jamie Cummins <Jamie.Cummins@UGent.be>
Sent: Friday, November 8, 2019 9:53 AM
To: Payne, Keith <bkpayne@email.unc.edu>
Cc: Sean Joseph Hughes <sean.hughes@ugent.be>; ian hussey <ianhusseysipod@gmail.com>
Subject: Re: Follow-up based on review of PSP-A-2019-0728

Dear Keith,

Thank you for your thoughtful and detailed comments to our initial email. Over the past few weeks we have talked amongst ourselves and really tried to grapple with your various points. Below we respond to each of them in turn:

I don't believe the new design eliminates the post-hoc confabulation account. As subjects are completing the AMP, they can form an impression of the target as pleasant or unpleasant before making either the awareness report or the target evaluation report. So the fact that they are reporting awareness before reporting the target evaluation doesn't mean that they are not reflecting on their evaluation of the target when making their report of awareness. And if they are reflecting on their evaluation of the target when reporting awareness, then there is the possibility to report post-hoc inferences about why they had that evaluation. For what it's worth, I would expect the awareness reports to be smaller in this design that your previous design, because in this design they are only making inferences about their subjective evaluation, whereas in the previous one they are reporting about both their subjective evaluation and the explicit report

of the evaluation. But in any case I think the standard explanation of misattribution + post-hoc inference could explain better-than-chance reports of awareness in this modified design.

1. If we understand this point correctly, you are suggesting that participants can form an impression of the target before they complete either the awareness rating or the evaluative rating task, and as such, it is possible that they can also reflect on that evaluation prior to completing either of the above two measures. As such, evaluative ratings might be compromised by the covert target evaluation that the person made earlier in the trial.

Upon reflection, it seems that the easiest way to control for this claim is to create a modified version of the AMP wherein the awareness question is presented *before the target ever appears on screen* (see the attached figure). In such a design, participants first see the prime, are then asked if the prime will influence their subsequent response to the target, and only then see the target stimulus which is subsequently evaluated. We believe that post-hoc inference could not play a role in such an experimental setup given that the influence question is asked before the target stimulus is even seen.

Before we collect data with such a procedure, we would like to ask you for your perspective on this study. Specifically, **do you think this new design eliminates the possibility of post-hoc confabulation**? If not, we would sincerely appreciate further clarification on this. Specifically, (a) in what way do you think confabulation could still play a role, and (b) is there an alternative design that would convince you that post-hoc confabulation was not playing a role in the AMP effects reported in our studies? If so, could you clarify what this might entail?

Using the reports of awareness to "postdict" priming effects in another AMP would still be problematic in this design. Based on the misattribution + post-hoc inference account, I would predict that subjects who show larger priming effects in one AMP should also report more awareness in the other AMP. So, we both expect a positive correlation. Why not just report the correlation? Using the "prediction" or "post-diction" language and conditioning on awareness reports when predicting AMP performance just brings in causal rhetoric when describing a correlational finding.

2. We definitely take the point that the language in the manuscript around the relationship between influence awareness and AMP effects needs to be sharpened up. Reviewer 1 also raised a similar concern, and we plan to be more specific on this front in our revision.

I don't think this design aligns with the logic of our skip study. Our point was about the relationship between awareness and control over subsequent responses. We expected that subjects might skip at rates better than chance, and hence reduce the priming effect when allowed to skip. But we predicted that they would not have such good awareness that they could eliminate the priming effect altogether. That is, if subjects had good enough awareness to allow perfect control over responses, we would not expect to see priming effects in a standard AMP. It turns out that in that study, they didn't reduce the priming effect at all. But I would not be surprised if (at least under some conditions) subjects show non-zero awareness effects in the skip condition based on post-hoc inferences. For example, suppose we had run another AMP, as in your studies, and selected subjects with the largest priming effects. I would expect these subjects to skip more often than those with small priming effects (because they can observe that their evaluations of the target usually line up with their evaluations of the primes). Still, I would not expect those subjects to be able to eliminate the priming effect by skipping, for the same reason described above.

3. Hopefully our above proposed design can now subvert the issues associated with post-hoc inference, and that we can better address the question of the role of influence awareness in the AMP. However, I think our findings from our extant studies (issues of confabulation aside) align with your position: although influence-awareness lead to much greater AMP effects, it did not remove the priming effect entirely. That said, our results clearly show that it did reduce it, contrary to what you say (i.e., not reducing priming effects at all). We also think that it's important to appreciate our section in the manuscript on structural validity which speaks to our related concerns here: while it appears that influence-unaware trials continue to produce priming effects (albeit of a reduced magnitude in our experiments, see above), the measurement properties of these trials seem to be problematic. Ie, although a priming effect is demonstrated, the measures does does not seem to function as a good measure of evaluations when using only these trials.

I wonder if we are working with different notions of what a misattribution means in this procedure. By misattribution, I mean that subjects cannot fully disentangle their evaluation of the prime from their evaluation of the target. So when a pleasant prime is presented and subjects evaluate the target positively, they know that the prime was positive, and they know that they responded positively. If you were to ask them if their response was due to (1) the prime valence (2) the target valence (3) both, they might well say both when they give prime-consistent responses. It's not as if a misattribution requires that they don't believe the prime had an influence. It's that once the prime and target are presented, their influences are now blurred together so that attempts to remove the prime effect from one's evaluation are unlikely to

succeed fully. A misattribution can co-exist with awareness that the prime influenced one's evaluation of the target.

4. In terms of discrepant definitions of misattribution, this could well be the case. Our argument within the general discussion of the paper was that if effects in the AMP are (primarily) the result of influence-aware trials, then this at minimum would require a reappraisal of the definition of misattribution. This was because, in our understanding, misattribution as a mechanism is considered to only operate under unintentional and unaware conditions. We arrived at this position primarily as a result of a quote from your original 2005 paper:

"We suspect that if participants recognized that their judgment on any given trial was being influenced by the prime, they would be able to correct by simply giving the opposite response. However, if they felt inclined to evaluate a pictograph positively because the pictograph actually seemed pleasant, then an adjustment would seem inappropriate".

Of course, we are wholly open to the possibility that we either mischaracterised this position, or that understandings of misattribution (both in general and specifically within the AMP) have evolved since the 2005 paper.

We'd like to be explicit more generally on the fact that our work intends to address misattribution only as a secondary issue, rather than as part of the focal implications of our work (and we intend to adjust the manuscript to explicate this more clearly). In general, our research questions are intended to most primarily address to (1) whether the AMP is implicit (in the sense of unaware), and (2) whether the AMP is a (structurally) valid measure of evaluations in all participants. Our discussion point on misattribution was mainly meant to highlight the potential implications of our findings for the mechanism under the assumption that unawareness and misattribution are related.

Overall, do you think our new proposed design (prime -> awareness question -> target -> evaluative response) can effectively avoid the confound of post-hoc confabulation, and allow us to more clearly investigate the role (if any) of influence awareness in AMP effects? If not, could you tell us what aspect of this new design you consider to be problematic?

Thank you again for engaging with us on this topic, and we look forward to hearing from you again.

All the best,

Jamie

Ian

Sean

On 15 Oct 2019, at 21:55, Payne, Keith <bkpayne@email.unc.edu> wrote:

Hi Jamie,

Thanks for your note. Here are my thoughts:

1. I don't believe the new design eliminates the post-hoc confabulation account. As subjects are completing the AMP, they can form an impression of the target as pleasant or unpleasant before making either the awareness report or the target evaluation report. So the fact that they are reporting awareness before reporting the target evaluation doesn't mean that they are not reflecting on their evaluation of the target when making their report of awareness. And if they are reflecting on their evaluation of the target when reporting awareness, then there is the possibility to report post-hoc inferences about why they had that evaluation. For what it's worth, I would expect the awareness reports to be smaller in this design that your previous design, because in this design they are only making inferences about their subjective evaluation, whereas in the previous one they are reporting about both their subjective evaluation and the explicit report of the evaluation. But in any case I think the standard explanation of misattribution + post-hoc inference could explain better-than-chance reports of awareness in this modified design.

2. Using the reports of awareness to "postdict" priming effects in another AMP would still be problematic in this design. Based on the misattribution + post-hoc inference account, I would predict that subjects who show larger priming effects in one AMP should also report more awareness in the other AMP. So, we both expect a positive correlation. Why not just report the correlation? Using the "prediction" or "post-diction" language and conditioning on awareness reports when predicting AMP performance just brings in causal rhetoric when describing a correlational finding.

3. I don't think this design aligns with the logic of our skip study. Our point was about the relationship between awareness and control over subsequent responses. We expected that subjects might skip at rates better than chance, and hence reduce the priming effect when allowed to skip. But we predicted that they would not have such good awareness that they could eliminate the priming effect altogether. That is, if subjects had good enough awareness to allow perfect control over responses, we would not expect to see priming effects in a standard AMP. It turns out that in that study, they didn't reduce the priming effect at all. But I would not be surprised if (at least under some conditions) subjects show non-zero awareness effects in the skip condition based on post-hoc inferences. For example, suppose we had run another AMP, as in your studies, and selected subjects with the largest priming effects. I would expect these subjects to skip more often than those with small priming effects (because they can observe that their evaluations of the target usually line up with their evaluations of the primes). Still, I would not expect those subjects to be able to eliminate the priming effect by skipping, for the same reason described above.

4. I wonder if we are working with different notions of what a misattribution means in this procedure. By misattribution, I mean that subjects cannot fully disentangle their evaluation of the prime from their evaluation of the target. So when a pleasant prime is presented and subjects evaluate the target positively, they know that the prime was positive, and they know that they responded positively. If you were to ask them if their response was due to (1) the prime valence (2) the target valence (3) both, they might well say both when they give prime-consistent responses. It's not as if a misattribution requires that they don't believe the prime had an influence. It's that once the prime and target are presented, their influences are now blurred together so that attempts to remove the prime effect from one's evaluation are unlikely

to succeed fully. A misattribution can co-exist with awareness that the prime influenced one's evaluation of the target.

I hope these thoughts are of some use. All the best,

Keith

Keith Payne, Ph.D.

Professor and Director of Graduate Studies

Department of Psychology and Neuroscience

From: Jamie Cummins <Jamie.Cummins@UGent.be>
Sent: Friday, October 4, 2019 10:21 AM
To: Payne, Keith <bkpayne@email.unc.edu>
Cc: Sean Joseph Hughes <sean.hughes@ugent.be>; ian hussey <ianhusseysipod@gmail.com>
Subject: Follow-up based on review of PSP-A-2019-0728

Dear Keith,

Many thanks for your recent review of our paper "The AMPeror's New Clothes: Performance in the Affect Misattribution Procedure is Mainly Driven by Awareness of Influence of the Primes" for JPSP. We were extremely grateful for your extensive, thoughtful, and comprehensive comments. We are currently busy designing one or more studies to address the concerns that you and the second reviewer highlighted.

To ensure that these new studies are of maximum utility, we would like to involve you in the design process and to include our correspondence in the preregisterations of any future studies. Our goal here is to only run those studies that directly address the issues you have raised. We would therefore appreciate your input on the logic of our arguments (*see below*), our characterisation of the existing literature, and how we can empirically address your concerns.

One of the main points in your review was that our various experiments (like most other published studies on this topic) failed to eliminate post-hoc confabulation as a potential explanation for AMP effects that are moderated by influence awareness. In other words, by asking participants about the influence of the prime on their target evaluations *immediately after* they had emitted those evaluations, any response to this question may still represent a post hoc confabulation on their part. You argued in your review that, if so, then our results are consistent with the misattribution account; i.e., "participants can observe their own behavior and notice if they are responding in prime-consistent ways. If so, they can report afterward that they were influenced by the prime".

We think a relatively simple modification to our existing paradigm would allow us to eliminate post-hoc confabulation as a potential explanation of our findings. Specifically, in our previous AMP experiments, participants were presented with a prime → blank screen → target → mask (which remained onscreen until an evaluative response was emitted). Only after emitting this evaluative response were they given the opportunity to emit an influence-awareness response (i.e., to indicate if the prime stimulus influenced their target evaluation). The aforementioned trial sequence can be altered in order to eliminate the possibility of post-hoc confabulation: the following trial sequence could instead be presented: prime → blank screen → target → mask and opportunity to emit an *influence-awareness* response → opportunity to emit target evaluation. That is, the order of presentation of the evaluative response and the influence-aware response are simply switched. In this way, one's influence-awareness response cannot be confabulated with one's target evaluation because the influence-awareness judgement is made*before* the evaluation itself is emitted.

If you agree with us, then we propose to run an experiment that is very similar to Experiment 2 in our manuscript but uses an IA-AMP which is modified as outlined above (see the attached Figure 1). As in Experiment 2, this would also examine whether the influence-awareness rate in the IA-AMP is *postdictive* of the absolute magnitude of the AMP effect on a previously completed standard AMP.

From our reading of Experiment 3 from Payne et al. (2013), this paradigm wholly aligns with the rationale which you used in the development of your 'skip' paradigm (i.e., participant reports of influence-awareness should be emitted before any evaluative response).

Would you agree with us that the aforementioned method effectively overcomes the issue of post-hoc confabulation? If not, could you tell us what, for you, is the feature that distinguishes between this proposed paradigm and that used in Experiment 3 of Payne et al. (2013), which you argued was effective in eliminating trials with prospective awareness?

We would appreciate your feedback on this revised version of our paradigm. If you disagree with our approach, we would also appreciate if you could tell us the precise issues that you see with it, so that we could further try to design a paradigm which can also overcome those potential issues. You can find the Inquisit script for the proposed paradigm (as well as the relevant additional experiment files) attached to this email. We also would like to thank you again for taking the time to review our manuscript and for providing critical and highly useful feedback.

Best regards,

Jamie Cummins

Ian Hussey

Sean Hughes