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 when preregistering those studies. We only want to run studies that directly speak to your questions and would therefore appreciate your input on the logic of our arguments, our characterisation of the existing literature, and how your concerns can be empirically tested.

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 emit those evaluations, any response to this question may still represent a post hoc confabulation on their part). You also argued 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”.

Before we can speak to whether influence-awareness is post-hoc or not, we would first like to have you clarify what you mean by an “evaluation” and when that evaluation is said to occur. Simply put, we need to be clear about *what* an event is before we can determine whether other events can, or do, occur before or after that event. We see two possibilities here.

First, one can define a “target evaluation” as the **overt** **response** that people make towards the target stimulus(i.e., the behavior of rating the target as positive or negative using the E and I keys). If you define target evaluation in this way then we think a relatively simple modification to our existing paradigm would allow us to control for the post-hoc confabulation argument you raised. Specifically, in our previous AMP experiments, participants were presented with a prime 🡪 blank screen 🡪 target 🡪 mask (until an evaluative response was emitted). Only after this overt evaluative response were they given the opportunity to emit an influence-awareness response (i.e., indicate if the prime influenced their target evaluation). If your position is that the target evaluation is the **overt evaluative response** then the trial sequence can be altered to eliminate a post-hoc confabulation account. That is, participants can be presented with the target 🡪 blank screen 🡪target 🡪 giving an opportunity to emit an influence-awareness response 🡪 and only then the opportunity to emit an evaluative response. In this way, the influence-awareness would not be confabulated after the target evaluation is made given that it is provided before the evaluation itself. As far as we understand, this is the position that resulted in the design of your third Experiment in Payne et al. (2013). If you agree, then we propose to run an experiment that is very similar to Experiment 2 in our manuscript but uses an IA-AMP that is modified as described above (see Figure 1). As in Experiment 2, this would 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.

*Mask & Evaluative Response*

*Prime*

*Blank Screen*

*Target*

*Prime Influence Question*

***Unpleasant***

***Pleasant***

***Unpleasant***

***Pleasant***

***Pleasant***

***Unpleasant***

***Pleasant***

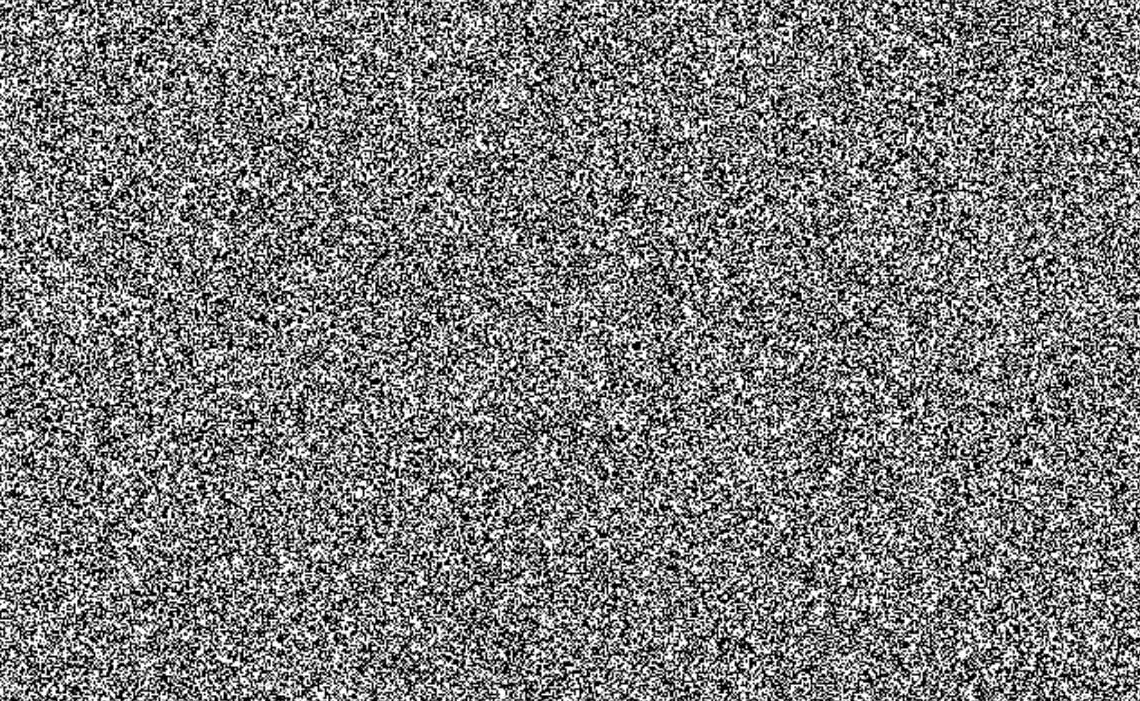
***Unpleasant***

***Unpleasant***

***Pleasant***

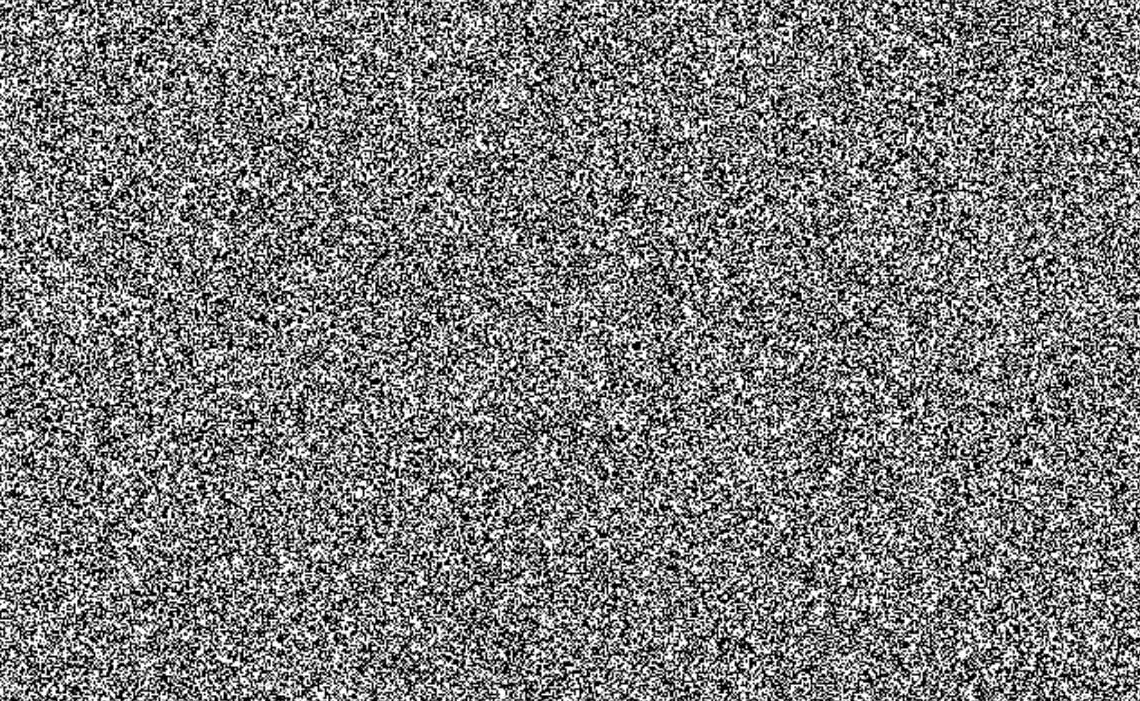


*Press spacebar if the picture will influence your response to the Chinese symbol*



*Figure 1*. Trial sequence for a typical trial in the proposed study.

Alternatively, if you define target evaluation as a **covert response** or as an **event that occurs at the mental level** (e.g., the formation of an association between positive valence and the target stimulus representation), then the above paradigm would not speak to the post-hoc confabulation account. When defined in this way the “target evaluation” would presumably occur prior to the overt response being emitted (i.e., people see the prime 🡪 see the target 🡪 form a mental association between the target and positive valence 🡪 which mediates their subsequent overt evaluative response). This argument is agnostic to exactly when the evaluation (in the sense of covert responses or mental events) occurs, other than that it necessarily occurs after the presentation of the prime *and* target, and before the overt response being emitted. On the one hand, we can still control for the post-hoc confabulation account here. Specifically, participants can be presented with the target 🡪 blank screen 🡪 giving an opportunity to emit an influence-awareness response 🡪 target 🡪 mask and opportunity to emit an evaluative response. In this way, influence-awareness would not be confabulated after the target evaluation is made given that it is provided before the evaluation itself. If you agree, then we propose to run an experiment that is very similar to Experiment 2 in our manuscript but uses an IA-AMP that is modified as described above (see Figure 2). As in Experiment 2, this would again 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.



***Pleasant***

***Unpleasant***

*Target*

*Prime Influence Question*

***Unpleasant***

***Pleasant***

***Unpleasant***

***Pleasant***

*Mask & Evaluative Response*

*Blank Screen*

*Prime*

***Unpleasant***

***Pleasant***

***Pleasant***

***Unpleasant***

*Press spacebar if the picture will influence your response to the Chinese symbol*

*Figure 2*. Trial sequence for a typical trial in the proposed study.

From our reading of the previous AMP literature and your own review, our understanding is that you endorse the first position (i.e., that target evaluation refers to the overt evaluative response emitted on a given trial). We arrived at this understanding for several reasons.

First, in Payne et al (2013) Experiment 3, you created an AMP variant in which participants were provided with an opportunity to “skip” rather than providing an evaluative response should they feel that their responses had been influenced by the prime. You report that the absence of significant differences in the mean AMP effect size between a standard AMP and this modified ‘skip’ AMP represented evidence that the AMP effect was indeed driven by influence-unaware responding. We see two possibilities here. On the one hand, if the first (overt) definition of ‘target evaluation’ is adopted then your argument in the 2013 paper holds: as the evaluation was emitted prior to the assessment of influence-awareness, this measure of influence-awareness was not post hoc, and thus the results can test (i.e., provide evidence for or against) the misattribution account. On the other hand, if one adopts the second (covert) definition of ‘target evaluation’ then the results you present Experiment 3 of your 2013 paper cannot test the misattribution account. Specifically, although participants who choose to ‘skip’ a given trial have not provided an overt evaluative response, this is uninformative as to whether or not they have emitted a covert evaluative response, or whether evaluation at the mental level has occurred. Given that you present the results of that experiment as representing evidence for the misattribution account, we have to conclude that you do endorse the first definition of ‘the evaluation’.

* If you stand over Paynes 2013, then you must accept that our proposed study would test the misattribution account.
* Conversely, if you don’t accept that our proposed study above would test the misattribution account, then you must explicitly acknowledge Payne 2013 as not testing the misattribution account either, and moreover are arguing that the misattribution account is untestable (in the sense that all assessments of influence-awareness are posthoc by definition, and therefore ‘accommodated by the account’).

At this point, what is the utility of the Covert response IA-AMP that Sean proposes (i.e, study 7)?

. One idea we have had involves taking further inspiration from the paradigm of Experiment 3 from Payne et al. (2013). Specifically, we had intended to use our IA-AMP paradigm, but to ask participants about whether their evaluation of the target will be influenced by the prime *before* any response is given (i.e., thus avoiding the confounding issue of retrospective confabulation). From our perspective, this approach overcomes the previous issues of retrospective self-report which you pointed out in our work, while simultaneously providing more information than the original 2013 paradigm (i.e., it allows for us to compare within-subject AMP effects based on trials which are prospectively described as influenced vs. uninfluenced).

Before running with this paradigm, we were hoping to get your opinion on its suitability to addressing questions relating to influence awareness in the AMP. Do you think that such a method successfully overcomes the issue of using post-hoc self-report measures? And more generally, do you think that this revised method could be useful to address whether influence awareness plays a role in effects which are seen in the AMP?

If you have the opportunity to provide us with any feedback on this revised version of our paradigm, we would greatly appreciate it. You can find the Inquisit script of the revised paradigm 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