**Supplementary Materials**

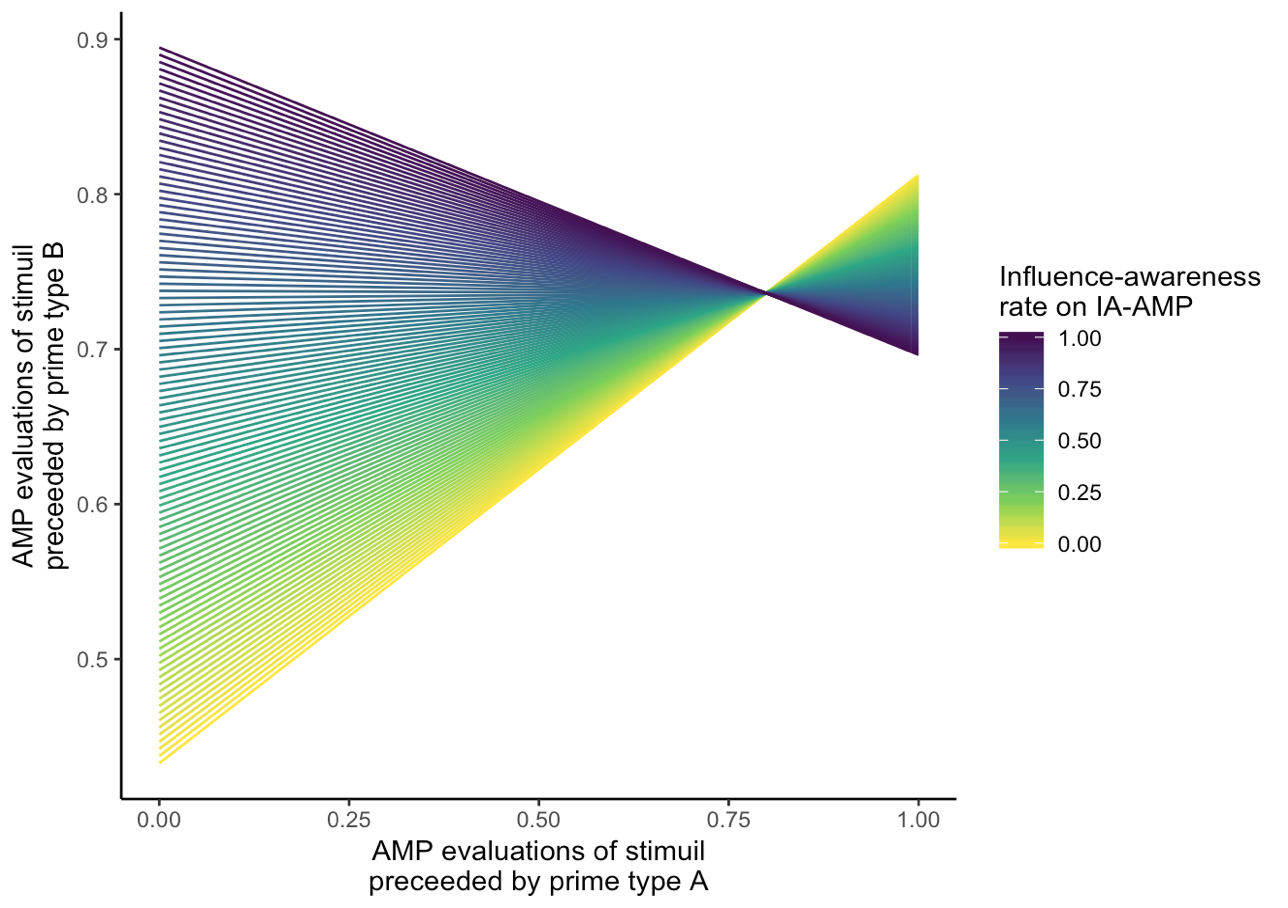
**Structural validity: Does the AMP measure**

Finally Lastly, we examined an aspect of the AMP’s structural validity. Whereas the previous analyses demonstrate that the AMP is not an *implicit* measure of evaluations, these analyses examine whether, aside from its implicitness, the AMP can be said to be a structurally sound *measure of evaluations* at all.

First, it should be noted that recent work has highlighted that a measure’s structural validity is vital to theoretical development which relies on the results of that measure, yet thorough consideration of structural validity is commonly overlooked [(Flake & Fried, 2019; Hussey & Hughes, 2018)](https://www.zotero.org/google-docs/?ruHERi). Throughout this article we have employed a common method of scoring the AMP: the difference score between evaluations of trials preceded by one prime type (e.g., positive stimuli) versus trials preceded by the other prime type (e.g., negative stimuli). From a measurement perspective, it is important to note that this (and other similar) AMP scoring and analytic methods imply a specific factor structure: two negatively correlated latent variables (i.e., evaluations of the two different prime types; see Rose, Wagner, Mayer, & Nagengast, 2019 for relevant discussion on the latent variable structures assumed by simple scoring metrics). If evaluations on the AMP are impacted of the prime stimuli - regardless of whether this impact occurs inside or outside of awareness - ratings of stimuli preceded by one prime versus the other must be negatively correlated between participants. For example, the more negatively participants rate negative-prime trials, the more positively they should rate positive-prime trials. The corollary of this simple point is that evaluations of the two trial types should not be positively correlated if the measure’s structural validity is sound. If positive correlations were to be found, this would represent evidence that evaluations are not being differentially impacted by the two classes of prime stimuli. We therefore assessed whether correlations between evaluations on the two prime type trials were positive or negative, as an important metric of the AMP’s structural validity. This analysis was not preregistered.

As in previous analyses, we examined moderation of this relationship by the influence-awareness rate observed in the IA-AMP. As noted in previous analyses, because the AMP was completed prior to the IA-AMP, assessing influence-awareness in the IA-AMP could not perturb the traditional AMP effect. As in the third meta-analysis, above, suitable data were therefore taken from experiments that included both an AMP and IA-AMP (i.e., Experiments 2, 3, and 5). A linear mixed-effects model was constructed using the lme4 R package [(Bates, Mächler, Bolker, & Walker, 2015)](https://www.zotero.org/google-docs/?bPjtuP), with mean ratings on prime B trials as the DV, and mean ratings of prime A trials and influence rate as the IVs. Experiment and AMP domain were entered as crossed random effects.

The key effect was the interaction effect. Results demonstrated that the correlation between ratings of the prime type A and B trials on the AMP was strongly moderated by participants’ influence-awareness rate in the later IA-AMP, *B* = 0.63, 95% CI [0.55, 0.71], β = 0.75, 95% CI [0.65, 0.85], *p* < .0000001. Specifically, participants’ ratings on the two prime type trials correlated negatively when their influence-awareness rates were high, indicating that primes influenced evaluations, and therefore structural validity in this subset of participants. However, correlations were negative when influence-awareness rates were low, providing direct evidence that primes did not influence evaluations in this subset of participants (see Figure 5). Given the distribution of influence-awareness rates, this suggests that he AMP demonstrates poor structural validity in the majority of participants. As such, where the previous analyses suggest that the AMP effect is not an *implicit* measure of evaluations, the current results additionally suggest that the AMP effect is not a structurally-valid measure of evaluations at all in a large subset of participants.



*Figure 5.* Evaluations of AMP trials preceded by the two prime types are correlated negatively when influence-awareness is high, but positively when influence-awareness is low. Positive correlations suggest poor structural validity.