Thank you very much for the positive comments regarding our manuscript and for the opportunity to submit a revised manuscript. We carefully considered the comments raised by the reviewers, and have edited the manuscript accordingly. Below we describe how we responded to each of the reviewers' points.

Reviewer 2 noted that when considering Figures 2 and 3, some parameter values used in the simulation are more plausible than others and that the most plausible values are those where true causal effects lead to results that would be flagged as spurious using the triangulation method. We very much agree with the reviewer, as this observation supports our primary argument. Our only issue in addressing this comment is that, as the reviewer hints, the ground truth is unknowable. This means that assessing plausibility in a defensible way can be challenging. To address this issue, we added a bit more description in the initial section describing our choice of parameter values to use in the simulation. Here we rely on empirical evidence about typical stability coefficients and lagged effects. We now highlight that certain values (i.e., very strong lagged effects) were chosen for demonstration purposes, despite the fact that they are not plausible. Please note, however, that the inclusion of the most implausible values does not help our argument: these values are those that are least problematic for Sorjonen et al.'s approach. So, our approach to interpreting these results is the more cautious one.

When describing the results themselves from Figures 2 and 3, we decided not to add argumentation about the plausibility of specific values. We did leave in our reference to the results of the Orth et al. (2022) review, as that empirical evidence is relevant for assessing the plausibility of different values. However, although we very much agree with the reviewer's assessment of the plausibility of different results (an assessment that supports our primary claim that the triangulation method is likely to detect effects as spurious even when they are not), we thought that additional speculation beyond what we provided would either take too much space to justify or (more problematically) rest on assumptions for which we do not have direct empirical evidence. Thus, we mainly left this section as it was. We did, however, briefly discuss this issue in the discussion section, highlighting that what empirical evidence does exist about typical stability coefficients and lagged effects suggests that the most plausible values are those where the triangulation test is most likely to fail.

The reviewer notes that without information about the "ground truth" regarding the size of autoregressive and cross-lagged effects, our statement that "the triangulation method does not add information about the likelihood that an effect is spurious" may not be correct. The reviewer did not specifically ask us to change the paper to address this comment, but we want to respond in this letter. Our analyses show that it is possible, under realistic data-generating models, for the triangulation method to detect an effect as spurious even when it is not. Even if we knew the ground truth about what effects are typical in the field, the triangulation test would not provide additional information in any specific case. In any individual case where the method was applied, it would be impossible to determine whether the effect was flagged as spurious because it was truly spurious or because the true pattern of correlations, stability coefficients, and lagged effects combined to create a false-positive when the triangulation test is used.

Reviewer 2 suggested that we modify Figure 1 to distinguish between "Factor X" and "Measure X". Currently, X is treated as a latent variable in Panel A and an observed variable in Panel B. If we were developing a similar example from scratch, we would have used something like the approach that the reviewer suggested; we agree that it would be clearer. The reason why we chose to retain the original is that our goal was to clarify the precise models tested in the paper to which we respond (i.e., Sorjonen, Arshamian, et al., 2024), something that the original authors did not do. Because of this, we wanted our figure to map as closely as possible to the example used in that original paper. We have reread the authors' example again, and we think that this figure (where X and Y are latent variables in the data-generating model, and then the same X variable is used as a predictor in a simple regression equation when testing the causal effect model) is the most accurate representation of the authors' description of their data-generating model and analysis. We considered adding "X1" as an observed variable that was predicted (without error) from the latent X variable, but this addition was not explicit in the original authors' paper, and we thought that this version did not map as closely to the original as we would have liked. Thus, we chose to retain our original figure.

Finally, the reviewer notes that a critical issue concerns the question of why initial levels of X and Y are correlated, as these reasons affect model choice. Again, the reviewer did not make specific suggestions about how to address this issue. We wholeheartedly agree with the point that the reviewer is making, but we also thought that addressing it would take us too far from the primary point of our paper, which is focused on the validity of the triangulation method. We do believe that our introductory material, where we address the challenges of using longitudinal data to address causal questions, hints at the issues that the reviewer raises. We also think, however, that the issue that the reviewer raises is a much bigger question about the types of models we should be using to assess causal effects with longitudinal data, and this question is, to us, outside of the scope of our paper.

Again, we thank you and the reviewers for their thoughtful comments on the paper. We hope that you find this revised version acceptable for publication in the European Journal of Personality.