Section names and levels: Should "Disturbance-Dependence Plots" and some other following sections be underneath "Our Approach"? Right now, they are at the same level as "Our Approach".

Yes, they should. This has been corrected.

p. 1: Instead of "visualizing statistical models is not so easy", maybe you mean "visualizing statistical fits of LVMs is not so easy"?

Yes. This was corrected.

Top p. 2: This paragraph currently talks about what is in the paper. I think the main contributions of the paper could be explicitly added here. I think the reviewers need to see that you are describing new LVM visualization approaches, illustrating the approaches, and providing software functionality for generally applying these approaches to other models/datasets.

Good suggestion. I modified to text to talk about how our approach is different, can detect things other strategies can’t detect, and is able to visualize just-identified/equivalent models.

p. 2: Instead of "To combat this", maybe "To address this issue"? Also, at the end of that paragraph, I wonder what the word "significant" means. I think you mean something like "useful", but others could read it as "statistically significant".

It looks like we removed this part of the paper in this revision.

Bottom of p. 2: Feel free to ignore this, but I would soften phrases like "suffer from a number of problems", in order to avoid arguments with reviewers on a point that is not really central to the paper. Like I would call out the problems but not say that a method is suffering from them. For "problems" here, you might include difficulty/availability of implementation, along with overwhelming amounts of numerical output that can be difficult to reconcile into an overall model assessment.

Good suggestions. Done!

Top p. 3: I don't think it is safe to assume that people know Anscombe's quartet, so it would be helpful to add a little detail to that sentence to describe what it is.

Done.

Figure on p. 4: In the text, it would help to make it clear that this figure shows two separate models/panels, as opposed to the entire figure being a two-factor model with uncorrelated latent variables.

Good suggestion. Done.

Bottom p. 5 (comment for the future, not for this paper): Is it correct that the non-unique factor score estimates are resolved by Bayesian models? I think the non-uniqueness just means that people are summarizing the posterior distribution in different ways, but Bayesian instead gives you the full posterior distribution.

Yeah, I suppose it’s a different problem in Bayesian. But, I think it’s an easier problem to resolve.

Top p. 6: At the start of the "Our Approach" section, it would be helpful to provide an overview about what visualizations flexplavaan provides that are not described by previous LVM work. This will help signal the unique contribution of the current paper.

Great suggestions. Done.

Hopper plots: It might be helpful to explicitly say that, as we go from the top of the plot to bottom, the correlations go from high to low. And it might be helpful to say what we should be looking for in this plot: that the top of the funnel is not too wide?

Good suggestion. We’re just looking for small values at the top of the hopper.

Trace plots: I think that the Bayesians have already taken this name for MCMC diagnostics. I think it is not a big deal to re-use it here, but there was also the funnel/hopper issue that you mention in the section before. More generally, I think it would help to highlight that the "x" and "y" variables are interchangeable here: the model itself is not making a statement about one observed variable causing another (at least, when the two observed variables come from the same latent), but it implies a correlational relationship between the observed variables that can be represented as a regression line.

I changed the name to “trail” plots. Also, I mentioned the x/y distinction is arbitrary.

Equations on p. 8: Maybe there are different notational systems, but I personally would not square the covariances. For example, I would write sigma(x,y) without a square. Then the covariance of x with itself, sigma(x,x), is also the variance and could be written as sigma^2(x).

Good point. This was changed.

Bottom p. 11: When you say "fit implied by the model", do you include latent variables in the prediction? In other words, does each observation get a unique prediction, or are you just subtracting the same mean off of everyone?

The latter.

Measurement plots, p. 13: For Psych Methods, I think more detail is needed about what is being shown in Figure 10. The phrase "puts all variables on a common scale" is a bit vague. I imagine the computation is similar to the trace plots, where you get model-implied covariances (for observed vs latent variable) and residual variances and then compute the slope and intercept.

Done.

Also, I couldn't tell why the underestimation is stronger for the Jedi lv... the right panel of Fig 10 looks pretty similar to the left panel.

That was a misinterpretation of the plots. My mistake. I fixed it.

Fig 11: Along with a loess line, showing a regression line through the points might reinforce that it is a linear relationship.

Good suggestion.

Fig 12: You might mention that this is a "product indicator" approach to handling the quadratic effect. See, for example Little et al 2006, On the Merits of Orthogonalizing Powered and Product Terms: Implications for Modeling Interactions Among Latent Variables.

Done.

Psych Methods typically values some code to be included with the paper. Say, the code that you actually used in the paper, or a brief document like "getting start with flexplavaan". I typically put a "computational details" section just before the references to provide details about where to access the code, so reviewers don't miss it.

Great idea!

One idea for a data application: it would be cool if the plots could be used to distinguish between two possible models like in Figure 2. That could really demonstrate the added value here.

Great suggestion. In addition to fitting simulated data for Figure 2, we have also added some other examples where we compare models.

For better or worse, I used my middle initial in my first couple papers, so I'd request my name be listed as "Edgar C. Merkle".