Memo for U2P

Your paper was well structured, and the experiment you did was clear. For the experiment itself, I could find few improvements. The analysis of two models and F-test for the comparison was very good. So please understand my feedbacks are more about technical issues than analytic issues.

In the first paragraph of the Method part, the audiences may be confused about the reason for grouping the participants several times. It may be better to explain here why the participants were separated and regrouped several times at the start of, or during the paragraph.

In the third paragraph of the Method part, I think we can remove the sentence 'Next, A brief comparison… in the model', the last sentence of the paragraph. The first sentence of the next paragraph says the exact same thing.

In paragraph 4 and 5 of the Method part, instead of just referencing 'In order to perform an F-test for comparison between two models, …', it would be better if 'I performed H0 = ~ versus HA = ~ because one model must be a sub-model of the other in order to perform an F-test.' It feels this information was provided after everything was finished. I assumed this F-test is what you used for 'the quality of the fit of the new model'. It may be better if you say you will use F-test for the analysis here. There is no problem with understanding the content of what you want to say, but I think the order of information provided in the paper is regrettable. Not only this part, but the explanation of the viable linear model part also seems not well-mixed with the contents that are explaining. If you mix the contents or put this information in front of the paragraph, it would be good.

Also, I think you did not concern the intended audience much for your paper overall. There are lots of explanations about the data and model in your paper, However, there are fewer explanations about the statistical methods you used for the analysis. I know it is awful to explain the whole thing, but it would be better to just explain what this is in short. For example, the intended audiences, undergrad/grad students who majored in psychology or high-school students, may find it difficult to what the F-test, degree of freedom and p-value are. Or they would not know the meaning of low in-sample mean square error means. If they do not have an idea of what they are, they cannot understand

where your results came from. If you add just one or two sentences to explain what those are, it would be good.

The caption of the figures is too long and provides too much information. Captions explain what the figures mean, not what we can get from these figures. At least the mean-square error should only be in the text, not in the caption. Also, I think it was a cool idea to make links to each figure and reference that you want to use but coloring them may be a little distracting. I think just using black letters would be fine.

One thing I thought about is the sample size. There are only 70 participants here. If you use the model $E[O|C, IT, G] = \beta_0 + \beta_1 C + \beta_2 IT + \beta_3 C * IT$, this means you would make two model for the case where C = 0 and for the case where C = 1 since C is the categorical variable. In other words, you split the data into two groups and made two models correspond to each group. If you add the variable G, you would like to give advantage or disadvantage for different genders. Then we know that the participants for the Easy Question group are about 30, which may give us a biased result. Even a smaller number of participants in the group if we use the graph for the analysis (there are only 11 participants in the Easy, Male group). We cannot quantitatively analysis about this low sample size problem, but I think it would be good to talk about this problem in the discussion section.

Little confused about what the first model and second model were. Not necessary but would be better if you clearly define the first model and second model in your analysis like 'I will call this model as the first/second model.'