Thank you for the revised version of this Stage 1 RR, and your careful responses to reviewer comments. I am sufficiently satisfied with these that I do not think it necessary to involve the reviewers again at this point. However, there are still some minor issues to be sorted out fully before IPA is issued for this study, and I list these below.

Response: Thank you, these additional questions are really helpful. We have provided responses below and changes in the main document.

1) There is still ambiguity about how the outcomes of the different hypothesis tests will be combined to inform conclusions with respect to the overarching theoretical question (LPP vs HPP). What follows may seem like nitpicking, but it is crucially important to be clear up front about how your conclusions will follow from your results.

At line 145 you state the overall rationale for the study as follows:

"If prior expectations are weaker in VR (LPP account), the magnitude of both the SWI and the MWI will be smaller in VR compared to the real world (see hypotheses H1A and H1B in table of questions). Additionally, the difference in peak grip force and load force rates between small and large objects (SWI), or more and less dense-looking objects (MWI), will be smaller in VR than in the real world (see hypotheses H2A and H2B in table of questions)."

What you seem to state here is that H1a AND H1b AND H2a AND H2b all test the same hypothesis (and the final column of your design table implies the same). If this is the case, then the implication would be that they must all have significant outcomes (in the same direction) for the hypothesis (LPP or HPP) to be supported. If so, then the hypothesis would be confirmed by the congruent conjunction of all four outcomes, and not by any other set of outcomes. This may or may not be what you mean to state, but it is what I would infer from what is written.

If this is not what you mean to state, then you will need to specify how you would interpret a set of results that provided partial support for one hypothesis (i.e. that had some outcomes significant but not others), or where different outcomes supported different hypotheses (LPP vs HPP). At present, it is not clear how you will draw conclusions across the pattern of results. This arises because all hypothesis tests are related back to the same global hypotheses.

Response: In the section of the paper that you reference we really meant that weaker influence of priors could lead to *any* of those effects, rather than necessarily *all* of them, so should have used 'OR' rather than 'AND'. We think that if prior expectations have relatively less influence any of those effects could be observed, but it is entirely possible (for either experimental or mechanistic reasons, e.g., measurement noise) that prior expectations could exert relatively less (or more) influence and not all effects emerge. And that an absence of an effect wouldn't (in isolation) automatically reject the LPP or HPP account.

As discussed in the response to the next comment, for the statistical analysis we will treat H1a, H1b, H2a, and H2b as separate hypotheses about the priors governing each of the perceptual and motor aspects of the SWI and MWI tasks. We will however use the combined pattern of results to determine whether there is overall evidence suggesting generally stronger or weaker influence of priors in VR. Our conclusions would be weighted by how many of the effects emerged – one statistically significant effect consistent with the LPP and three null effects would be considered support for this explanation, albeit fairly weak. Whereas four statistically significant effects would be considered strong evidence. By contrast if there was a mixture of findings, with statistically significant effects in both directions, we would infer that the results were inconclusive about a broader explanation using the LPP or HPP and that neither could be accepted.

We have added this information to the paper [line 145].

2) This uncertain status is compounded by the design table, in which you state:

"NOTE: these hypotheses for the SWI and MWI tasks are being treated as individual hypotheses that are related to the same question¬¬, rather than employing a disjunctive or conjunctive logic (Rubin, 2021).

There is some linguistic ambiguity here. Either they are individual (i.e. separable) hypotheses, in which case they lead to conclusions on separate questions, or they relate to the same question (in which case they are two tests of the same hypothesis). The word 'relate' may be intended to mean only that there is a thematic relationship between the separable questions (e.g. LPP vs HPP for SWI; LPP vs HPP for MWI), but you need to be clear about whether the tests can lead to separate conclusions or not.

Response: Thanks, yes we really meant that they are individual hypotheses (and are being treated as such statistically) but there is a thematic relationship to the LPP and HPP explanations. We are treating LPP and HPP as broader hypotheses about VR, made up of each of the slightly different questions (priors about object size and materials for perception and for action). Changes made in the table of questions.

3) In the design table, in several places, you make statements about the conclusions that will follow from non-significant outcomes, e.g. "No statistically significant difference between conditions would indicate no difference in strength of prior expectations."; "No statistically significant difference in pGFRdiff scores would indicate no difference in strength of prior expectations." Etc. In NHST, a failure to find a significant difference does not allow one to accept the null hypothesis (only to fail to reject it), and so you should remove these statements.

Response: Thanks, yes we really meant 'would indicate no EVIDENCE FOR a difference in strength of prior expectations' so have removed these as suggested.

4) You are very clear that the main hypothesis tests will not be meaningful if manipulation checks are not passed. I would advise you to state which checks apply to

which tests (e.g. failing a manipulation check for the MWI would presumably not stop you testing hypotheses relating to the SWI). Also, if a manipulation check is failed, so that a hypothesis test is deemed uninformative, does this mean that you will not run that hypothesis test? If so, state this. If you will still run it, then state why (given that it will be uninformative).

Response: Yes, if it was not met, the corresponding hypothesis tests would not be run, so we have specified in each case which hypothesis the check corresponds to and that the test would not be run.

5) In the Abstract, you have the statement: "hypothesis posits increased reliance on predictions relative to current sensory information due to sensory uncertainty." I think you should probably delete the "due to sensory uncertainty", which wrongly implies that the cause of a relative change in weighting could be known in this experiment.

Response: Thanks, changed.

6) "windsorised" >> "winsorised"

Response: Thanks, changed [line 300].