Minor Revision

We are almost there. I just have one query. In the discussion, you say "The Bayesian rmANOVA generated the highest BF10 for the model that included only the TMS site factor (BF10 = 3.46), showing that the observed data are better represented by considering the ipsilateral and contralateral differences. The model including only the TMS timing factor produced a very low BF10 (BF10 = .05)". As the first reviewer said, in effect, the last test includes variance from the first factor in the error variance; but we have just concluded that the first factor involved more than error variance. So you address this here (and for the other analysis where this comes up), by averaging over models with and without the factor (if you stick with this, make this clearer in the legends for tables 5 and 7). OK, but then you still need conclusions to follow from the same rules of inference you have been using, i.e. 1/3 < B < 3 is non-evidential - as the reviewer indicated. (I feel slightly happier just basing assumptions on a model with main effects included at once, otherwise one includes in the average over models, the case just mentioned, where the error variance is wrongly estimated. But as this is exploratory, I leave that detail lto you.) To address this point, you could just report the one analysis that best estimates error variance for each test, or average as you do - but in any case make sure the conclusions follow the rules of inference you have already decided on.

Dear Prof. Dienes,

We thank you for clarifying this point. We have updated our Stage 2 accordingly and we provide a revised version.

Specifically, we have rephrased our conclusions for our exploratory analyses in pg. 33 to make them less strong, and to reflect our decision threshold as follows:

"[…] In detail, the inclusion of the TMS site factor resulted in the highest BFincl (BFincl = 23.01), but including the TMS timing factor in the model, resulted in lower posterior odds, which even though they did not reach our decision threshold (1/3 < BF10 < 3), they point towards evidence against (i.e., BFincl < 1) a model including the TMS timing factor (BFincl = .73) or an interaction of TMS site and TMS timing (BFincl = .34). The results of the Bayesian rmANOVA inform us that, in line with the registered analyses of Experiment 1, a TMS site (ipsilateral vs contralateral) effect is evident and that no timing differences are unlikely, as reflected by the evidence against a model which solely includes timing or a model that includes an interaction with timing, and the low posterior odds for an averaged model which includes the TMS factor."

Also, we did similar revisions for the analysis in Experiment 2, in pg. 34:

"[…] Most models including solely TMS timing, or TMS timing interactions resulted in low BF_{incl} (all BF_{incl} < .22; see Table 7 for details), with only the TMS site by TMS timing interaction failing to reach our decision threshold (1/3 < BF_{10} < 3), with evidence favoring no such interaction (BF_{incl} = .36). The results of the Bayesian rmANOVA are analogous to those

registered for Experiment 2, where both a TMS condition (sham vs real) and TMS site (ipsilateral vs contralateral) effects were found, but differences across timings are unlikely."

Lastly, we included the phrase *"BF_{incl} is calculated as the likelihood ratio representing the change from prior odds to posterior odds for each factor in the model averaged by all the models that include each factor"* in the legends of Table 5 (pg. 60) and Table 7 (pg. 62).