

The authors have responded to my previous comments and made some changes in response but also rebutted quite a few of my comments, which to some extent of course is their prerogative and I respect their autonomy as researchers to make ultimate decisions. I do have some additional suggestions including a critical one related to the assumed effect sizes that were used to estimate sample sizes needed.

Line 156 “beyond mere detection of acoustic periodicities” This is odd to say when followed by cues that include tempo and rhythm, which have obvious periodicities that are relevant to the metre. Even timbre could be presented with periodic fluctuations (e.g., alternating plucking and bowing of a violin) that emphasize a particular metre.

Line 177 “a more direct effect of movement-related processes on metre” How does this model entail a more direct effect of motor mechanisms compared to predictive-coding and neural resonance theory. They all seem to have similarly direct interactions of sensory and motor processing, and as I pointed out in the last set of reviews and now acknowledged later in this section, it’s not clear how distinct any of these theories really are from each other.

Line 183 “More radically, the action simulation for auditory prediction” What’s so radical about this theory?

Re: studies by Philips-Silver et al., the descriptions “body movement coordinated with a rhythmic pattern” and “the metre the individual had previously moved to” make it sound like the participants were purposely moving to the rhythms, whereas at least in some of their studies, participants were passively moved by an experimenter, implicating the vestibular system more than the volitional motor system. Please clarify this.

Line 206 “more direct methods” As I said before, what’s so direct about measuring brain activity and movement? If you mean this is a more direct measure of metre perception, how could they be more direct that actually measuring *perception*? You can see brain activity such as SSEPs even in participants that are not attending or perhaps even asleep so clearly the interpretation of such activity as perception-related is not direct. To be clear, I’m not saying that the proposed measurements aren’t informative, just that they aren’t more direct or “better” than perceptual measures, i.e., again there is no gold standard as the term “direct” implies. So just maybe don’t use the term direct, and instead say that convergent evidence is needed across different kinds of measurements to show how movement can shape metre processing.

Line 423 The term “statistical learning” is still used once.

Participants: you won’t exclude people with psychiatric or neurological impairments such as autism, schizophrenia, ADHD etc. that are associated with sensory-motor and temporal processing abnormalities?

Line 446. What does “(French and English excluded)” mean? Does that mean they can’t speak those languages or it doesn’t count as one of the languages from the African countries so they would have to speak English and an indigenous language?

For the power analysis, some of the effect sizes used, especially for the interactions, seem very large, e.g., $f=.89$ ($f=.40$ is already considered a large effect according to common heuristics). This seems unrealistic and based on prior experiments that themselves were probably underpowered and therefore led to imprecise estimates of effect size. Specifically, Chemin et al. only used 14 participants in each of the experiments and only Experiment 1 showed a significant interaction effect but not Experiment 2, with very different estimates of effect size in those two experiments, which was glossed over on page 9 when the study was discussed in this proposal. The movement condition manipulation is a type of biasing of perception and there is very good reason to believe these kinds of effects will be quite small based on other studies besides Chemin et al. Specifically, SSEPs in both vision and hearing are mostly stimulus driven with small effects of attention and perception-related effects.