

Review Reports

(doi.org/10.31234/osf.io/xky4j)

Summary

The present study performed conceptual replications examining which information (visual or auditory) dominates in evaluating performance in music; importantly, the present study used classical piano competitions as like the previous studies and expanded the same paradigm for the Tsugaru shamisen (i.e., a traditional Japanese folk musical instrument) competitions. As a result, the present study found significant interactions between domain (visual vs. auditory) and variance in quality (1st and 2nd place vs. 1st and low-placing performers) in both instruments. However, the trends were slightly different between the instruments. In the low variance, visuals dominated the judgment of the piano performance, while this fashion was not found in that of the Tsugaru shamisen one. Moreover, in the high variance, auditorys dominated the judgment of the Tsugaru shamisen performance, while this fashion was not found in that of the piano one. They discussed their findings along with the original and other replication studies (including the recently published paper of Wilbiks and Yi (2022)).

#General comments

I would apologize for the delay. The authors completed a good work and I'm happy to review this full paper. This 2nd-stage manuscript followed the protocol and the authors discuss their results appropriately. Thus, there is little problem. The following comments may help to brush up the present study.

#Specific comments

- 2nd paragraph in the "2.5 Power analysis" section: "... re-analysis of Mehr et al.'s data using using the parametric t-tests" -> "... re-analysis of Mehr et al.'s data using the parametric t-tests" (FYI, there are other errors of this kind)
- It is unclear what "H4, H5 and H6" mean, although I guess that H1-3 are for the piano condition while H4-6 are for the Tsugaru-shamisen one. Please clarify these.
- It is better to keep the significant figures of the statistics consistent within the manuscript.
- It is difficult to interpret the graphs of the relative effects (Fig4b and 4c). The relative effect values in the text do not seem to match the values in the graph; for example, they reported "relative effect = .69" in text for the visual domain in the low variance for the piano condition, while the relative effects seem to be under

.625. Why? Perhaps, is the "relative effect" different between in-text and in-graph? Anyway, it might be better to added some explanations.

-The visualization of the accuracy of each pair (i.e., Fig 6) is highly helpful. If my understanding is correct, the left-most pair (i.e., M.F. vs D.L.) has the opposite tendency to the results of the low variance for the piano condition in the confirmatory analysis. Like this, one of the Tsugaru-shamisen pairs (i.e., Y.W. vs N.K.) has a different trend from the results of the confirmatory analysis. It might be beneficial to discuss what could trigger these differences; of course, this might stem from just noise.

-On p.23, they stated "we suspect that the different cultural backgrounds of participants may have played a stronger role". It should be better to suppress the tone of this. At this time, as the authors mentioned, there is little evidence supporting this claim, and it is quite possible that the failure of the replication is due to the slight differences in experimental design.

-The authors proposed that visual salience might play a role in the participants' prediction of competition winners, in particular, in the case of the piano. This is an interesting idea, and the attempt to capture the mechanism of this phenomenon from the framework of information processing is commendable. For that reason, it should be desirable to discuss this issue in more detail (for example, how visual salience is involved, and how this salience is used in making judgments, etc).