

# Understanding the role of visual and auditory information in evaluating musical performance

**Yuki Yamada** based on reviews by David Hughes and Kyoshiro Sasaki

A recommendation of:

**STAGE 1**

## Open Access

Sight vs. sound in the judgment of music performance: Cross-cultural evidence from classical piano and Tsugaru shamisen competitions

Gakuto Chiba, Yuto Ozaki, Shinya Fujii, Patrick E. Savage (2021), *Psyarxiv*, xky4j, stage 1 preregistration, in principle acceptance of version 5 by Peer Community in Registered Reports <https://doi.org/10.31234/osf.io/xky4j>

Published: 28 December 2021

Copyright: This work is licensed under the Creative Commons Attribution-NoDerivatives 4.0 International License. To view a copy of this license, visit <http://creativecommons.org/licenses/by-nd/4.0/>

Submitted: 24 September 2021, Recommended: 28 December 2021

### Cite this recommendation as:

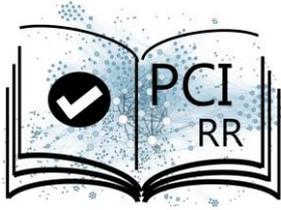
Yuki Yamada (2021) Understanding the role of visual and auditory information in evaluating musical performance. *Peer Community in Registered Reports*, 100003. <https://doi.org/10.24072/pci.rr.100003>

## Recommendation

In this Stage 1 Registered Report, Chiba and colleagues (2021) aim to investigate how people use information from visual and auditory modalities when evaluating musical performances. Previous studies, mainly using Western music, have reported a visual dominance, but this has not yet been clearly and consistently reported. Thus, the authors propose to evaluate both the reproducibility and generalizability of the previous findings by conducting a replication study using the methodology of the previous studies and by introducing a new experimental condition in which the Tsugaru-shamisen, a unique Japanese musical instrument, is also performed. This study could represent an important turning point in the research context of performance evaluation and would be of considerable value.

This manuscript was peer-reviewed by two experts in scientific methodology and Japanese traditional music, respectively, and during the two-round peer-review process they made a number of important points, but eventually awarded the manuscript a highly positive response. I am therefore pleased to recommend that this Stage 1 Registered Report meets our Stage 1 criteria and is worthy of in-principle acceptance. I look forward to seeing the results and discussion reported in Stage 2, with the expectation that the experiment conducted by the authors will be in strict accordance with this protocol.

\*The following is a very minor comment, which I hope the authors will find helpful in the future. Of course, this is not related to hypothesis construction and does not require



revision: The "Blind Audition" study cited in the introduction is very impactful, but has recently been called into question, so I am at least a little cautious when citing this study. This article may be useful. <https://www.wsj.com/articles/blind-spots-in-the-blind-audition-study-11571599303>

**URL to the preregistered Stage 1 protocol:** <https://osf.io/ry2b6>

**Level of bias control achieved:** **Level 6.** *No part of the data or evidence that will be used to answer the research question yet exists and no part will be generated until after IPA.*

**List of eligible PCI RR-friendly journals:**

- [F1000Research](#)
- [Peer Community Journal](#)
- [PeerJ](#)
- [Royal Society Open Science](#)
- [Swiss Psychology Open](#)

## References

1. Chiba G, Ozaki Y, Fujii S, Savage PE (2021) Sight vs. sound in the judgment of music performance: Cross-cultural evidence from classical piano and Tsugaru shamisen competitions [Stage 1 Registered Report]. Psyarxiv, xky4j, stage 1 preregistration, in-principle acceptance of version 5 by Peer Community in Registered Reports. <https://doi.org/10.31234/osf.io/xky4j><https://doi.org/10.17605/OSF.IO/R2B6>

## Reviews

Toggle reviews

***Reviewed by [Kyoshiro Sasaki](#), 28 Dec 2021***

I'm happy that the authors have addressed my concerns. I recommend that this protocol should be approved.

---

## ***Evaluation round #1***

DOI or URL of the report: <https://doi.org/10.31234/osf.io/xky4j>

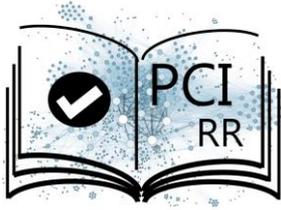
## ***Author's Reply, None***

[Download author's reply](#)[Download tracked changes file](#)

***Decision by [Yuki Yamada](#), 03 Nov 2021***

I appreciate your submission to PCI RR. As you can see, we were able to receive peer reviews from two relevant researchers: one is a cognitive psychologist who is very familiar with registered reports. The other is a widely experienced expert in Japanese historical music. Before introducing the individual peer review results, I would like to inform you that this manuscript requires a major revision before it can be recommended. The reasons for this are as follows.

The former reviewer seems to acknowledge the potential significance of this work, but also points out several major issues. These may be summarized in the appropriateness of the experimental design and the justifiability of the sample size design. In particular, if this study is truly considering that historical factors (I



am not sure if that is the suitable terminology) are related to the audience's performance evaluation, it should be specifically stated as a hypothesis, as pointed out by the reviewer, and the experimental design should be capable of examining it. That is, it is worth considering using a method that can detect the effects of knowledge about the historical background of the Tsugaru shamisen and about traditional performers (not the recent popular ones), and adding other popular music as an additional control condition.

The latter reviewer appreciates the article very much, but says that some technical expressions should be annotated. In fact, readers who read registered reports and are familiar with hypothesis-testing studies will have no difficulty in understanding the meaning of statistics and methodological abbreviations. However, this study has a very unique focus of research subject (i.e., Tsugaru shamisen), and the readership may be much broader than the authors envisioned. Therefore, it would be beneficial for the social impact of this study to supplement the descriptions with points even if the authors might feel they are redundant in writing usual manuscripts. However, in my opinion, adding detailed explanations of statistics in the text may reduce the readability for experts, so I thought it would be a good idea to use footnotes.

Thus, I am looking forward to receiving this manuscript again, which has been greatly improved by the review comments of both reviewers.

**Reviewed by *Kyoshiro Sasaki*, 02 Nov 2021**

[Download the review](#)

**Reviewed by *David Hughes*, 26 Oct 2021**

Chiba, Fujii & Savage: Sight vs. sound in the judgment of music performance  
review by David W. Hughes

This is a fascinating and valuable article. I myself have performed Tsugaru shamisen in Japan before audiences, and I have also (despite trying to avoid it!) been a judge at a few folk song contests, some of which have included Tsugaru shamisen performers.

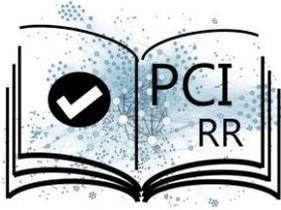
I'd never thought about whether sight (i.e. the appearance of the performers, their facial expressions, movement, clothing etc) influenced my judgment of performers, sometimes overcoming sonal differences between contestants (not only shamisen players but also singers). This article finally makes me think about it – about music as a “multimodal phenomenon”, to quote the authors.

I have toured the UK (as lecturer, co-performer etc) with two different folk music groups from Japan which included high-level Tsugaru shamisen players. Though these performances were concerts, not contests, I was definitely aware that bodily movement, facial expression and other visible elements impacted on the audience, and I even advised the performers of this. In fact, the performers were already aware of the importance of visibility. So indeed this article is important for bringing this to my conscious attention, and to realise that surely sight is indeed competing with sound for many judges and audiences.

This article should definitely be published and publicized. But for readers like me, more detailed explanation of some terms and concepts will be needed.

This article is quite technical, sometimes using terms that elude me a bit. On p.1, line 27 of the abstract, “ $d = 0.4$ ” confused me. I searched the internet for this phrase, and it seems to link with “Cohen's  $d$ ”, though even then I couldn't completely grasp it. This article does explain it somewhat, as an “effect size”, but perhaps a bit more explanation is needed for non-specialist readers. Also in section 2.3, the terms “p-value”, “t-test” and “alpha level” need more explanation, at least for people like me! Also, “GC” in line 188 puzzled me, but I presume it refers to the co-author Gakuto Chiba.

They have also cited many writings that I, as a “normal” ethnomusicologist, have never read or even heard of. This is useful, in that those technical writings are surely important to the more scientific readers at whom the article is clearly aimed. Much of the analysis and discussion in this article will indeed appeal to and



interest people like me – as I noted above, it certainly made me think more clearly about the influence of visuals in music performance. But this article will also reach out to scholars focusing on sound perception and “cross-modal” analyses.

They also cite a range of writings about Western classical music, noting that scholars have often pursued “the role of visuals and sound” (p.3). Thus considering the sonic-visual differences in perception of other musical genres broadens such analysis in a valuable way.

One of their predictions (p.4, H1) is that “visuals will dominate the judgment” of the upper ranks of a Tsugaru shamisen competition, when the sonic performances are very close in quality. I’d never thought of that, but in fact I agree.

Various small changes are needed to help non-specialist readers. For example, on p.4, line 119, they write that they selected “brief 6s excerpts” for one test. I’ve finally realized that 6s = 6 seconds! But this was only made clear to me in line 327. They should change this to “brief 6-second excerpts”.

Still, the method for the test described there – having different people judge a performance in different ways, by only its audio, only its visual, or both together (audiovisual) – is excellent and indeed focusses on the theme of this article. And then comparing the participants’ judgments with the actual outcome of those performers in a contest helps us understand how audio and visual judgments can differ greatly.

Overall, despite some terminology eluding me, I truly look forward to the results of the full testing they will conduct, again focusing on the different perceptions by their test participants of audio, visual and audio-visual versions of performances. Thus I support their plans 100%. This article simply needs a bit more clarity in some places (mostly mentioned above) to make things clearer and easier for readers like me who are unlikely to read all the relevant publications in their bibliography.

CONCLUSION: Yes, this article deserves full support and publication.

Dr David W. Hughes (dh6@soas.ac.uk), 26 October 2021