

Hearing, seeing, liking: The effects of audio-visual listening conditions on perceptual ratings

Gerardo Lopez[†]

Leigh VanHandel²

¹ Michigan State University, East Lansing, MI, USA

² University of British Columbia, Vancouver, British Columbia, CA

[†] Corresponding author: spotgerry@gmail.com

Published 16 December 2021; <https://doi.org/10.18061/FDMC.2021.0019>

Author video presentation and/or other conference material: <https://doi.org/10.17605/OSF.IO/BY2HW>

Abstract

This study investigates whether listening to an excerpt of music in audio-only or audio-video contexts affects perceptual ratings, and what the relationship of the context presentation is to perceptual ratings. Participants listened to music in audio-only and audio-video contexts and completed a survey with self-reported ratings for familiarity, emotional and mental engagement, and likability. Results suggest there was no effect of audio-only or audio-visual presentation on perceptual ratings. This conclusion is similar to work done by Huang and Krumhansl (2011). Both emotional engagement and mental engagement had strong correlations with likability ratings. Emotional ratings had a slightly larger effect than mental ratings in the audio-only condition, but results are mixed in the video condition. This suggests a need for closer examination of the emotional and mental rating categories and their relationship to likability.

KEYWORDS: *perception, audio-visual modalities, emotional engagement, mental engagement, atonal/post-tonal*

Introduction

This study investigated whether listening to an excerpt of atonal/post-tonal music through audio-only or audio-visual contexts affect participants perceptual ratings, and if so, how the perceptual ratings relate to the presentation context.

Prior research in this area includes studies by Behen (1990) and Davidson (1993), both looking at effects of performance on expressivity or other evaluatory measures. Both found that auditory perception could be influenced by visual stimuli, specifically with musical performers. Huang and Krumhansl (2011) focused on potential modality effects (audio-only vs. audiovisual) on performance evaluation with a limited stimuli sample of three excerpts, of which only one was atonal/post-tonal; they found no significant differences attributable to modality of presentation.

One critical distinction between this study and previous work (including Behne, 1990; Davidson, 1993;

Huang and Krumhansl, 2011; Vuoskoski, Thompson, Spence and Clarke, 2016; Griffiths and Reay, 2018) is the prior studies' focus on performance ratings. In Huang and Krumhansl (2011), for example, they asked questions like "How appropriate is the performer's expressed emotion?" and "How well does the performance maintain your interest?", whereas the current study asked "How emotionally engaging did you find the musical excerpt?" and "How likeable did you find the musical excerpt?". In addition, this study is focusing on the effects of modality in both tonal and atonal/post-tonal repertoire.

Method

Participants

There were 31 participants (female = 17, male = 14). The age range was 18-46, with the average being 26 years old (SD = 7.986). The majority of participants were undergraduate or graduate students with 6-9 years of musical instruction and 2-3 years of music theory instruction. Participants were recruited via convenience sampling from online forums and were not compensated for their participation.

Design

The experiment was administered through a Qualtrics survey. The presentation of stimuli was randomized, each with the same set of four ratings questions regarding familiarity, emotional engagement, mental engagement, and likability (Figure 1). No specific definition for these terms was provided other than the familiarity rating requesting their level of familiarity with the excerpt before the study, in an effort to prevent higher ratings later in the repeated measures design.

Procedures

Participants heard and rated all of the stimuli, which consisted of 6 musical excerpts presented under two conditions, audio-only or audio-visual, resulting in 12



individual stimuli total (Table 1). The audio for the audio-only condition was sourced from the audio-visual stimuli and was not altered in any way. Participants heard between 30-40 seconds of the musical excerpt, from the beginning of the piece until an appropriate moment in the music (end of a phrase, rest, etc.). Afterwards, they were asked to self-report responses to the survey questions on a 1-6 Likert scale.

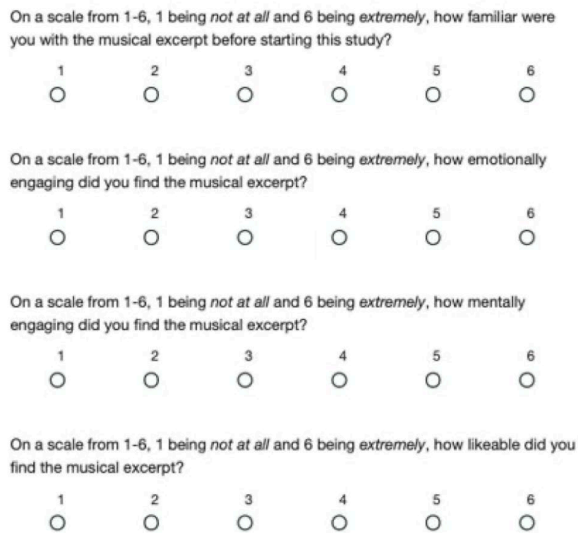


Figure 1: Survey questions used for all stimuli presentations (audio-only and audio-visual).

Stimuli

Excerpts were taken from six string quartets, including two tonal excerpts (Haydn and Beethoven) and four atonal/post-tonal works (Berg, Webern, Ligeti, and Carter). The excerpts were selected based on the availability of video recordings each featuring a professional string quartet that did not differ greatly in terms of performance setting, and that were filmed in a way that presented a realistic audience perspective as closely as possible.

Results

There was no significant difference in overall ratings for familiarity, emotional engagement, mental engagement, or likability across the audio-only and audio-visual conditions (Figure 2). For the individual excerpts, the only significant difference for the ratings across the audio-only and audio-visual conditions was the familiarity rating for the Ligeti excerpt, which saw a significant ($p = .003$) increase in familiarity in the audio-

visual presentation. All other familiarity ratings across conditions remained consistent.

Table 1: Stimuli information.

Composer	Piece	Period	Performer
Haydn, Joseph	Quartet in C Major, Op. 20/2, Mvt. I	Classical/ 1772	Quatuor Ebène
Beethoven, Ludwig van	Quartet in F Major, No. 11 "Serioso," Op. 95, Mvt. I	Classical-Romantic/ 1810-1811	Zemlinsky Quartet
Webern, Anton	Fünf Sätze (Five Movements), Op. 5, Mvt. I	Atonal/ 1909	Telegraph Quartet
Berg, Alban	String Quartet, Op. 3, Mvt. I	Atonal/ 1910	Telegraph Quartet
Ligeti, György	String Quartet No. 1 "Métamorphoses nocturnes," Mvt. I	Post-Tonal/ 1953-1954	Altius Quartet
Carter, Elliot	String Quartet No. 5, Mvt. I	Post-Tonal/ 1995	Pacifica Quartet

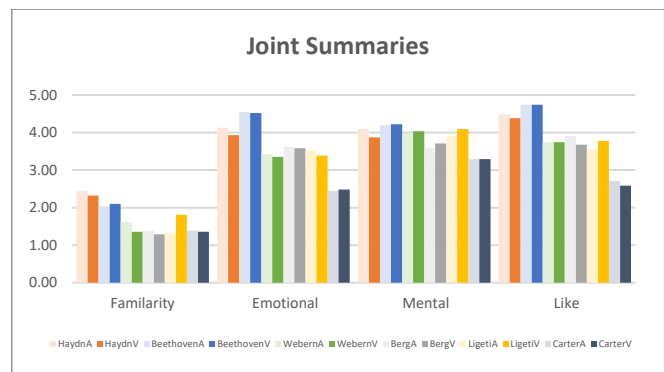


Figure 2: Chart with response averages for all excerpts across audio-only and audio-visual contexts. The “A” at the end of composers’ names indicates the audio-only context (lighter shading). The “V” at the end of composers’ names indicated the audio-visual context (darker shading).

Likability ratings were strongly correlated with emotional engagement ($r = .830$ for audio-only; $r = .808$ for audio-visual) and mental engagement ($r = .752$ for audio-only; $r = .733$ for audio-visual). Familiarity was also significantly correlated, but was weaker than emotional and mental engagement ($r = .3$).

One concern that arose from the general correlations was that the tonal excerpts were rated higher than the atonal/post-tonal excerpts, and that the higher ratings for tonal excerpts may be influencing the correlations. Additional post-hoc correlational analysis was conducted with excerpt groups to determine whether

there was a difference in style across the tonal vs. atonal/post-tonal conditions. The tonal condition combined Haydn and Beethoven (*HB*), atonal/post-tonal combined results for Webern, Berg, and Ligeti (*WBL*), and the results for Carter were their own category (*C*), since that excerpt’s ratings across mental/emotional engagement and likability were significantly lower relative to the other excerpts. Table 2 provides the results of the ratings from these excerpt groups, with further commentary in the discussion.

Table 2: Results from excerpt group correlations.

	Audio-Only			Audio-Visual		
	HB	WBL	C	HB	WBL	C
Emotional	.745	.805	.844	.741	.757	.832
Mental	.722	.757	.767	.757	.756	.723

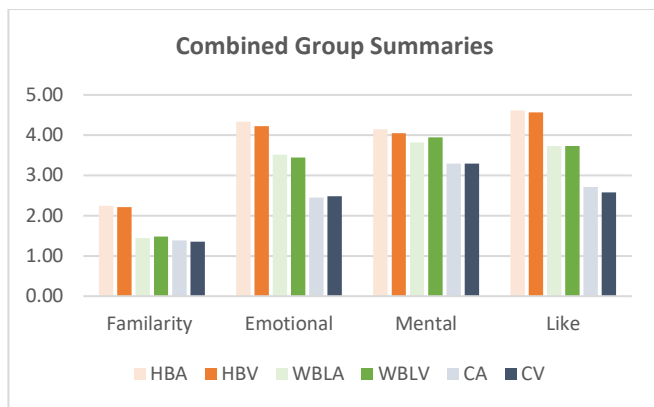


Figure 3: Chart with response averages for combined excerpt groups across audio-only and audio-visual contexts. HBA – Haydn and Beethoven ratings under the audio-only context; HBV – Haydn and Beethoven ratings under the audio-visual context. WBLA – Webern, Berg, and Ligeti ratings under the audio-only context; WBLV – Webern, Berg, and Ligeti ratings under the audio-visual context. CA – Carter ratings under the audio-only context; CV – Carter ratings under the audio-visual context.

Discussion

While there were no significant effects of likability or emotional and mental engagement across the audio-only and audio-visual conditions, the relationship between the likability and emotional/mental engagement ratings themselves is more varied.

From the ratings for all excerpts across the audio-only and audio-visual conditions, the likability and emotional engagement ratings had a slightly higher correlation than likability and mental ratings, though both correlations were very strong. When divided into the tonal vs. atonal/post-tonal excerpt groups, the high correlation between likability and emotional engagement remained in the audio-only condition, with emotional engagement ratings again being stronger than mental engagement ratings, but in the audio-visual condition things were not as clear. For the Haydn and Beethoven (*HB*) group the correlation between likability and mental engagement ratings was slightly higher than for likability and emotional engagement ratings; for the Webern, Berg, and Ligeti (*WBL*) group the correlations of likability and mental/emotional engagement were extremely close; the Carter (*C*) sustained a higher correlation between likability and emotional engagement ratings than for likability and mental engagement ratings. This suggests that there may be a need for future research to examine the relationship between likability and emotional and mental engagement more closely to clarify that relationship.

Additional analysis was conducted as an exploratory look at response differences between participants with more musical training and those with less musical training. Preliminary results suggest no overall differences, with the notable exception for mental engagement across audio-only and audio-visual contexts; participants with more musical training self-reported significantly lower mental engagement ratings than less trained participants. Future designs of this experiment will develop more systematic parameters for studying the differences of emotional and mental engagement in highly vs. less trained participants.

Some limitations of the current study include a relatively small sampling of only six stimuli, with an exposure to each stimulus for only 30-40 secs. In future experiments the number of stimuli would be increased to include more repertoire and potentially different exposure times. Another limitation is the different performance conditions, particularly in the audio-visual stimuli; ideally the recordings would be made by the same string quartet being recorded under the same performance conditions for all of the excerpts. Future studies may also investigate additional influences on the rating tasks, as well as differing representations of emotional and mental engagement. Such examples would include an expansion of the rating scale and the inclusion of more specific wording for the emotional engagement question to prevent participants from

equating only positively valenced emotions with being emotionally engaged. Additional questions on listener bias to atonal/post-tonal works can be included to control for its effect on ratings of engagement and likability.

One motivation for this study was the pedagogical question of whether presenting audio-visual recordings of performances of atonal/post-tonal works in a classroom setting would increase student engagement and likability over audio-only presentations. While the answer seems like a definite no, the results have shed light on a potentially much more complicated listener relationship warranting further investigation.

Conclusion

Overall, there was no significant difference between ratings for likability, emotional engagement, or mental engagement for performances presented in audio-only and audio-visual conditions. Some complex interactions between likability ratings and the emotional/mental engagement ratings indicate the need for further exploration of how genre and musical training influence these ratings.

References

- Behne, K. -E. (1990). "Blicken Sie auf die Pianisten?!" Zur bildbeeinflussten Beurteilung von Klaviermusik im Fernsehen [Do you watch the pianists? On visually influenced judgements of piano music]. *Medienpsychologie*, 2(2), 115–131.
- Davidson, J. (1993). Visual perception of performance manner in the movements of solo musicians. *Psychology of Music*, 21(2), 103–113. <https://doi.org/10.1177/030573569302100201>
- Griffiths, N., & Reay, J. (2018). The relative importance of aural and visual information in the evaluation by musicians and non-musicians of classical music performance. *Music Perception*, 364–375. <https://doi.org/10.1525/mp.2018.35.3.364>
- Huang, J., & Krumhansl, C. L. (2011). What does seeing the performer add? It depends on musical style, amount of stage behavior, and audience expertise. *Musicae Scientiae*, 15(3), 343–364. <https://doi.org/10.1177/1029864911414172>
- Vuoskoski, J. K., Thompson, M. R., Spence, C., & Clarke, E. F. (2016). Interaction of sight and sound in the perception and experience of musical performance. *Music Perception: An Interdisciplinary Journal*, 33(4), 457–471. <https://doi.org/10.1525/mp.2016.33.4.457>