VIBROTACTILE FEEDBACK ENHANCES PERCEIVED AROUSAL AND LISTENING EXPERIENCE IN MUSIC

Hanna Järveläinen
Zurich University of the Arts
Institute for Computer Music and
Sound Technology ICST
`hanna.jarvelainen@zhdk.ch`

Eric Larrieux
Zurich University of the Arts
Institute for Computer Music and
Sound Technology ICST
`eric.larrieux@zhdk.ch`

ABSTRACT

In this study, we measured the effects of vibrotactile feedback on the perception of music performance. Vibration signals were produced by transducers under a tabletop or under a chair and played simultaneously with an audio recording of solo cello music. Vibration types were either the signal recorded from the front plate of the cello simultaneously with the audio recording or white noise following the recorded amplitude. Perceived arousal was measured continuously from N=30 participants. In comparison to non-vibrating control conditions, especially sound-matching vibrations enhanced perceived arousal significantly. Increased amplitude of vibrotactile feedback had a positive but small effect on perceived arousal. In a post-experiment interview, participants described a higher sense of presence and embodiment with sound-matching vibrations.

1. INTRODUCTION

Vibrotactile enhancement of the auditory musical experience is an active research topic both in listening and interaction tasks due to the close relation of the senses of hearing and touch. Despite their individual spectro-temporal characteristics, auditory and vibrotactile perception show many similarities in pitch and intensity discrimination, energy integration over time, and masking effects [1]. Audio-tactile integration and interaction mechanisms are complex including effects in frequency, intensity, and spatio-temporal perception [2, 3].

Previous research in the field focuses on beat, meter, and rhythmic perception and implied positive effects on movement and appreciation [4–7]. Sensory substitution systems such as the Emoti-Chair [8] have been developed to make musical experiences more accessible to the hearing-impaired. Merchel and Altinsoy showed increased overall quality of the concert experience through a series of psychoacoustic studies involving various modifications of the vibration signal, such as low-pass filtering and compressed dynamic range [9]. Recent evidence underlines the importance of time-alignment and intensity congruence [10].

In active interaction, added vibrotactile feedback has enhanced enjoyment, perceived quality of the instrument or interface, and user experience, even if actual pitch selection or timing performance tasks have shown little or no effect [11–14].

Most of the previous research assessed experience after a listening or a playing period, however. The present setup measures continuous perceived arousal during listening. Along with valence, arousal is one of the two dimensions describing emotional perception of music in the circumplex model [15, 16]. Furthermore, as intensity changes over time are known to correlate highly with perceived arousal [17], we hypothesized that congruent vibrotactile feedback might contribute to perceived intensity changes through audio-tactile loudness integration and thus manifest as increased variability in arousal profiles.

We set out to measure continuous perceived arousal in music under varying types and intensities of added vibrotactile feedback, applying functional methods previously used in studies on audio-visual cross-modal interactions [18–21].

2. STIMULI

The auditory signals were two two-minute excerpts from pieces for solo cello. The cello was chosen because its pitch range matches the human vibrotactile perception range. The algorithmically generated compositions had comparable dynamic and tempo profiles as well as pitch range, style and set of playing techniques. To prevent participants from responding from memory at repeated hearing, the style was atonal contemporary, a genre unlikely to be easily memorized [1]. The excerpts were played on an acoustic cello by a professional cellist.

Audio recordings of the material were made from a distance of ca 2.5 meters using a Neumann 184 KM Stereo microphone set. The auditory recordings were played back through closed-back Sennheiser HD280 Pro headphones. Output level of the auditory feedback was set manually to a fixed and comfortable level. Closed-back heaphones were necessary yet sufficient to mask the sound of the transducers producing vibrotactile feedback.

---

[1] Quick tutorial on atonal music for the interested reader: https://youtu.be/gzodB0Sp6ZI

Figure 1. Transducers under the chair (top) and under the tabletop (bottom).

The respective vibrotactile signals were recorded simultaneously with the audio recording from the front plate of the cello ca 4 cm below the bridge using a SCHERTLER DYN-C Electrodynamic Cello Transducer. Two types of vibration signals were rendered and combined with the respective auditory signal: the recorded signal (hereafter "Signal" condition) and white noise following the amplitude of the recorded signal ("Noise" condition). The vibration signals were neither compressed in dynamic nor frequency range. Vibrotactile output was produced by two CLARKSYNTHESIS TST7239 Silver Tactile Sound Transducers, driven through an Audio DTA-1 amplifier and Fireface UCX II Audio Interface. One transducer was attached under the chair and one under the tabletop as seen in Fig. 1. Figure 2 shows example power spectral densities for vibration signals measured from the tabletop for Signal (tone E3 at 164 Hz) and Noise vibrations.

Two amplitude levels were used for vibrotactile feedback in the final experiment (see Section 3.1). First, a level was found that was clearly perceived by the authors. The other amplitude was then set 6 dB lower. It was furthermore expected that the same transducer output would produce stronger vibrations in the chair than in the more massive table. This was indeed the case, as seen in Fig. 3 showing example acceleration signals (Z-axis RMS dB re: 1 μm/s^2) measured by a vibrometer for High-amplitude Noise vibrations in the centre of the expected contact areas.

3. EXPERIMENT

3.1 Design

Perceived arousal was measured continuously under constant auditory and varied vibrotactile feedback manipulated according to two crossed variables: Vibration type [Signal, Noise, No vibration] and Vibration amplitude [High, Low, 0 (No vibration)]. Separate measurements were carried out for vibrations in the Table and in the Chair using sound samples 1 and 2, respectively, to reduce repetition. The resulting factor combinations, summarized in Table 1, were played back in random order.

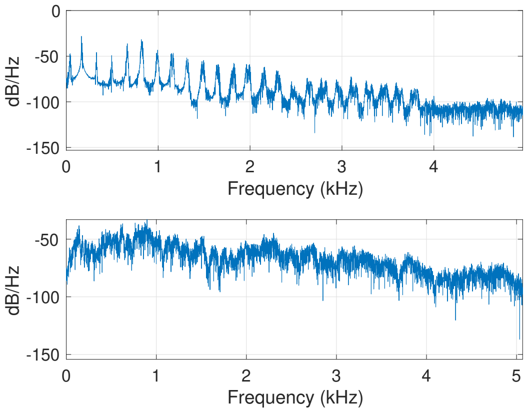

Figure 2. Power spectral density estimates for the vibration acceleration signals measured from the tabletop, top: Signal, bottom: Noise.

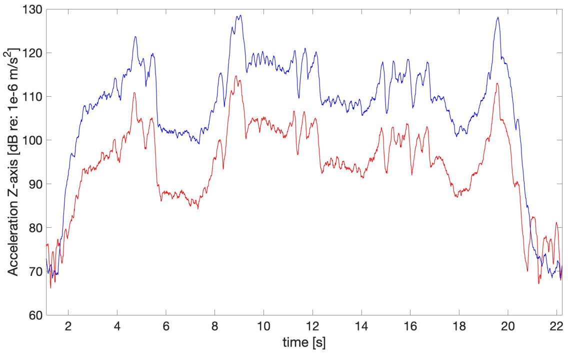

Figure 3. Initial 20 seconds of the High-amplitude Noise vibration signals (dB RMS acceleration in the Z-axis) measured from the Table (red) and the Chair (blue).

3.2 Participants and Procedure

Perceived arousal was defined as a subjective continuum varying from passive, sleepy, and low-energy for very low arousal to active, awake, and high-energy for very high arousal. Participants rated perceived arousal continuously during listening by controlling a continuous slider by means of a foot pedal. The slider end points were 0 (very low arousal) and 1 (very high arousal). Visual feedback was given of the current slider value. Sitting on the chair, participants kept their hands and forearms on the tabletop as seen in Fig. 4.

The measurement was carried out using custom-made software written in Max for controlling playback in Reaper and recording the responses. Changes in the slider value were recorded at 5-ms intervals and read into a file using time

Vibration type	Amplitude	Location	Audio
Noise	High	Table	Sample 1
Noise	Low (High -6dB)	Table	Sample 1
Signal	High	Table	Sample 1
Signal	Low	Table	Sample 1
No vibration			Sample 1
Noise	High	Chair	Sample 2
Noise	Low	Chair	Sample 2
Signal	High	Chair	Sample 2
Signal	Low	Chair	Sample 2
No vibration			Sample 2

Table 1. Summary of the test stimuli.

Figure 4. Experimental setup and user interface.

stamps synchronized with the respective sound file. Total listening duration was 20 minutes with short breaks at request.

The N=30 participants (average age 24.8 years, circa 50% males/females) were university students. Nineteen studied music or were musically trained. The rest had no significant musical training.

3.3 Results

3.3.1 Ratings

Raw data, after Loess-smoothing, is presented in Fig. 5 for trials with High-amplitude Noise and Signal vibrations versus no vibrations. Both Table and Chair vibrations seemingly enhanced the mean perceived arousal respective the non-vibration condition. The enhancement seems to manifest both as a constant shift and an increased range of variation. The effect seems stronger for Signal than Noise vibrations. In the Table data, the difference between Noise and Signal vibrations seems smaller and the peak responses seem slightly lower than in the Chair.

The musical excerpts, the dataset, and the analysis code are available in our repository [22].

3.3.2 Statistical Model

In order to make statistical inferences from the data, a statistical model was fit to predict the time-varying perceived arousal curve from scalar and functional predictors [23]. For this analysis, the data were registered at common 250-ms time intervals. Separate but identical models were fit to the Table and Chair datasets as follows:

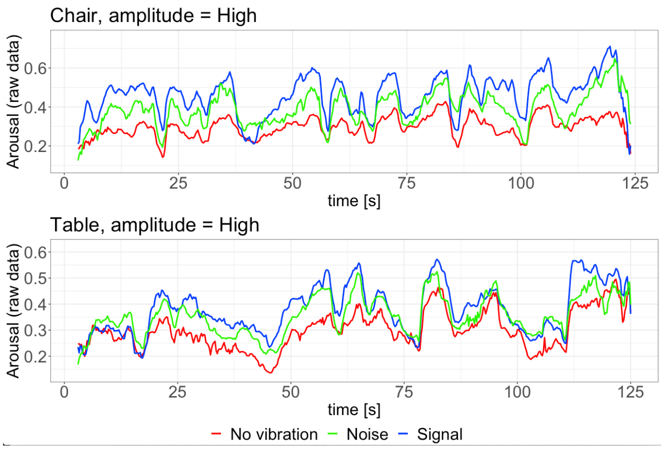

Figure 5. Mean arousal ratings for High-amplitude vibrations at Chair and Table (raw data).

$$\begin{aligned}\text{Arousal}(t) \sim & f(t) + \\ & \text{f.vibration}(t) + \text{vibration} + \\ & \text{f.amplitude}(t) + \text{amplitude}\end{aligned} \tag{1}$$

where

- Arousal(t) is functional perceived arousal
- f(t) is the general time-varying component due to musical content
- f.vibration(t) and f.amplitude(t) are time-varying effects of vibration type and amplitude
- vibration and amplitude are respective constant (scalar) effects

The model above is an additive functional regression model containing no interactions between the two predictors, vibration type and amplitude. The effects were estimated by approximate Bayesian inference in R [24] by the INLA method that uses integrated nested Laplace approximation, a fast alternative to Markov chain Monte Carlo sampling [25, 26]. A 2nd-order random walk model was specified in order to produce smooth curves.

Fitted arousal curves (posterior expected values from the model) and respective 95% credible interval [2] bands are presented in Figures 6 and 7. Estimated scalar effects, additive to the grand mean given by the Intercept representing the non-vibrating case, are presented in Table 2. Fig. 6 highlights the effect of amplitude separately for Noise and Signal vibrations. Ratings are interpreted to differ credibly between two conditions when their 95% CI bands do not overlap. Regardless of vibration type and location (Table/Chair), a common finding is that perceived arousal is credibly enhanced by added vibrotactile feedback. The effects are generally stronger for the Chair vibrations than

---

[2] Credible intervals (CI) are the Bayesian counterpart to confidence intervals in frequentist analysis

for Table vibrations. The effect of amplitude is also more notable, even though the CI bands for High and Low amplitudes overlap much of the time.

Figure 7 presents the differences between vibration types at High amplitude. Again for Chair vibrations, differences are credible and we conclude that the effect is strongest for sound-matching vibrations. For Table vibrations, the difference in favour of Signal vibrations is consistent but smaller so that the CI bands overlap much of the time.

Finally, Figure 8 shows the time-varying effects of vibration type and amplitude, in addition to the constant shifts in Table 2. The profiles show the time periods of their strongest impact. Similarly to the constant shifts, the time-variant components show larger differences between Signal and Noise vibration than High and Low amplitude, and altogether weaker effects in the Table. In the Chair, Signal vibrations boost perceived arousal especially during the first and the last 30 seconds. In contrast, the effect of High amplitude is strongest in the middle part. The differences might be explained through long tones in the opening and final passages which, according to many participants, brought the strongest benefits of Signal vibrations (see Sec. 3.3.3). Faster passages with mainly short bowing might then benefit from higher vibration amplitude and even Noise vibrations, as seen in the middle part of the rating profiles in the Chair.

Effect name	Chair	Table
No vibration (Intercept)	0.20	0.20
Vibration Noise	0.07	0.10
Vibration Signal	0.13	0.10
Amplitude Low	0.10	0.06
Amplitude High	0.13	0.07

Table 2. Estimated constant effects of vibration and amplitude.

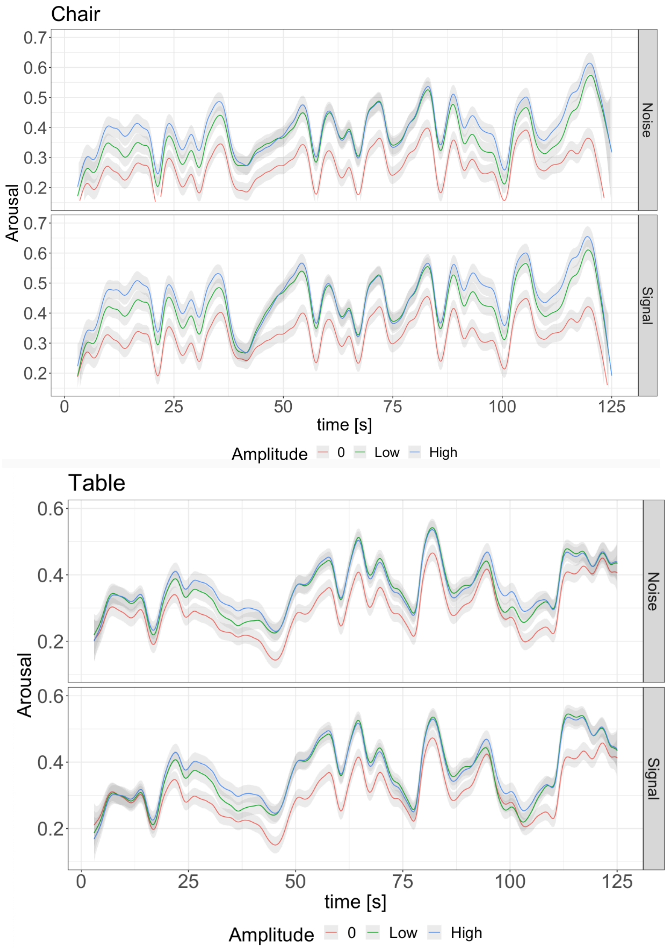

Figure 6. Estimated mean arousal curves with 95% CI bands for Noise (top panel) and Signal (bottom) vibrations.

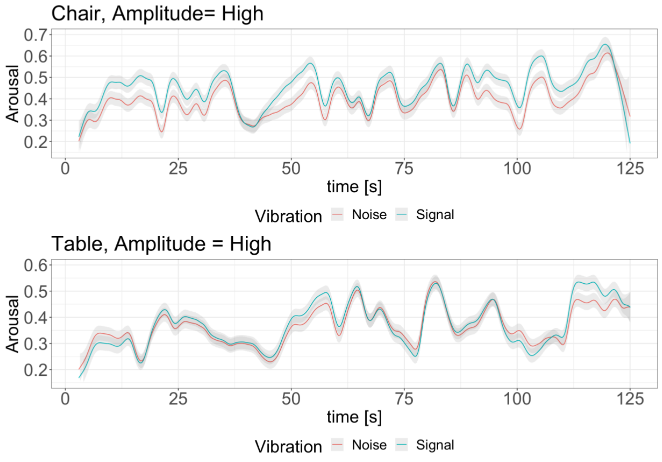

Figure 7. Comparison of Noise and Signal vibrations at Amplitude = High for Chair and Table locations.

3.3.3 *Interviews*

A short qualitative interview took place after participants finished the rating experiment. They answered the following two questions in their own words: 1) Did the vibrations always match equally well with the respective sound? 2) In case you noticed differences, please think back to the best-matching case and describe you experience (compared to the non-vibrating case)?

All participants noticed differences in the vibration patterns and reported the experience with their perceived best-matching combinations as positive in comparison to no vibrations. The three most frequent responses were increased sense of presence or embodiment, for example as if being in the same room or touching the instrument (21/30 participants), increased attention to the music and level of interest (13/30), and increased pleasantness of the experience (5/30). Two participants reported increased visual imagination of the playing gestures and one reported higher emotional connection. Note that these posterior responses cannot be tracked to the specific factor combinations. A frequent comment was, however, that long tones produced the largest benefits of vibrations matching the sound. In accordance with the difference in measured accelerations in favour of the chair (see Section 2), several participants felt the chair vibrations better, while only one participant felt vibrations stronger on the table.

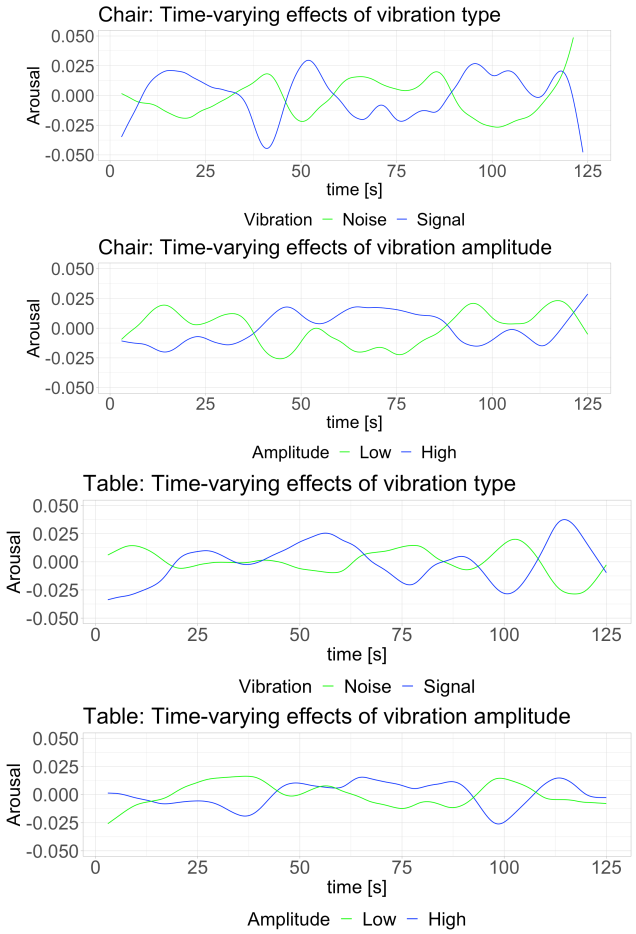

Figure 8. Estimated time-varying effects of vibration type and amplitude for Noise and Signal vibrations.

4. DISCUSSION

These results show that vibrotactile feedback enhances perceived arousal especially when the vibrations match the auditory feedback. The present study demonstrates the effect during listening; previous research has reported increased overall quality of the concert experience in presence of vibrotactile feedback, measured after listening [9]. The finding that sound-matching vibrations produce maximal enhancement is in line with observed benefits of vibrotactile feedback in active tasks with musical interfaces [11, 14]. Given that loudness is a known factor of perceived arousal, the effect might be caused by both increased average perceived loudness and and increased loudness variation resulting from multisensory integration [27]. Noise vibrations following the auditory amplitude produced less enhancement, however, even though such combinations should be rather similar to sound-matching vibrations in terms of loudness and synchronicity. There may be various reasons for why the combination with noise vibrations was less constructive. Recent literature supports the possibility that tactile noise could mask the auditory signal, at least on a cortical level [28], and tactile distractors are known to influence judgments of auditory intensity when the auditory modality is attended [2].

The effect of vibration amplitude was less prominent. Differences between High and Low amplitude were smaller than differences between Low and No vibrations, but both Low- and High-amplitude vibrations still enhanced perceived arousal credibly in comparison to no vibrations. A similarly weak but positive effect of amplitude was reported in a recent study [6].

In the present study the dynamic and frequency ranges of the vibration signals were not compressed. Previous studies have used various mappings from the auditory to the vibrotactile signal in order to make the vibrations perceivable inspite of differences between auditory and vibrotactile sensitivity [1]. Our hypothesis was that sound-matching vibrations might perhaps not be perceived all the time but that audiotactile integration would still cause increased perceived changes over time and thus enhance audiotactile perceived arousal. This was indeed observed in the measured profiles.

The generally weaker effects observed in the Table data than in the Chair can be attributed to higher mass of the table and the resulting weaker vibrations as Fig. 3 demonstrates. This difference was expected and accepted, as the auditory noise from the transducer would have become audible had we driven it at much higher power. A future study should perhaps focus on one contact location with maximal control of vibration amplitude, as such an extensive dataset is until today missing. Further uncertainty in vibration characterization results from participants' individual loads on the chair and the tabletop and their varying vibrotactile sensitivity. Such questions can hopefully be addressed in future research through new techniques to characterize haptic devices [29].

A detailed musical and acoustic analysis of the stimuli was not yet undertaken. A rough comparison was made between the musical content and time-varying effects of vibration type after several participants reported a very positive match between long notes and Signal vibrations. The time-varying components seem to reflect this but prompt a more thorough analysis.

The post-test interviews were short and entirely qualitative. Detailed questions in the middle of the session were not possible as participants were not informed about the varying vibration profiles during the experiment. They were told to pay attention to both sound and touch and instructed to maintain the listening position. The very positive results from the interviews were however clear and will motivate future research on vibrotactile enhancement of telematic performance or virtual sound environments.

The cello was chosen for this study specifically because its low pitch range matches human vibrotactile sensitivity which peaks at circa 250 Hz [30]. The present evidence cannot therefore be generalized to higher-pitched instruments such as violins. Moreover, the present case specifically concerns a near-field recording and vibrations recorded from the front-plate; it is uncertain if the effect would be present with recordings made at longer distance in a concert hall, even though low-frequency vibrations should carry to the audience.

5. CONCLUSIONS

Vibrotactile feedback had a positive effect on continuous perceived arousal in music perception, especially when the vibrations matched the sound. Effect of higher vibration amplitude was positive but less prominent. Vibrotactile enhancement was observed both in constant shifts in mean perceived arousal and increased perceived changes. These and other recent results exploring the important role of multisensory cues in music perception can contribute to emerging fields of music presentation such as telematic performance, immersive art, and virtual reality.

Acknowledgments

Thank you to Philippe Kocher and Jan Losos for generating and performing the stimuli, respectively.

6. REFERENCES

[1] S. Merchel and M. E. Altinsoy, “Psychophysical comparison of the auditory and tactile perception: a survey,” *J. Multimodal User Interfaces*, vol. 14, no. 3, pp. 271–283, sep 2020.

[2] J. M. Yau, A. I. Weber, and S. J. Bensmaia, “Separate mechanisms for audio-tactile pitch and loudness interactions.” *Front. Psychol.*, vol. 1, no. October, p. 160, jan 2010. [Online]. Available: http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=3157934{&}tool=pmcentrez{&}rendertype=abstract

[3] V. Occelli, “Assessing audiotactile interactions: Spatiotemporal factors and role of visual experience,” Ph.D. dissertation, 2010. [Online]. Available: https://core.ac.uk/download/pdf/35316637.pdf

[4] R. Brochard, P. Touzalin, O. Després, and A. Dufour, “Evidence of beat perception via purely tactile stimulation,” *Brain Res.*, vol. 1223, pp. 59–64, aug 2008.

[5] J. Huang, D. Gamble, K. Sarnlertsophon, X. Wang, and S. Hsiao, “Feeling Music: Integration of Auditory and Tactile Inputs in Musical Meter Perception,” *PLoS One*, vol. 7, no. 10, p. e48496, oct 2012. [Online]. Available: https://journals.plos.org/plosone/article?id=10.1371/journal.pone.0048496

[6] M. J. Hove, S. A. Martinez, and J. Stupacher, “Feel the Bass: Music Presented to Tactile and Auditory Modalities Increases Aesthetic Appreciation and Body Movement,” *J. Exp. Psychol. Gen.*, vol. 149, no. 6, pp. 1137–1147, 2020. [Online]. Available: /record/2019-66260-001

[7] C. Bernard, J. Monnoyer, M. Wiertlewski, and S. Ystad, “Rhythm perception is shared between audio and haptics,” *Sci. Reports 2022 121*, vol. 12, no. 1, pp. 1–12, mar 2022. [Online]. Available: https://www.nature.com/articles/s41598-022-08152-w

[8] M. Karam, C. Branje, G. Nespoli, N. Thompson, F. A. Russo, and D. I. Fels, “The emoti-chair,” in *CHI*, Atlanta, Georgia, 2010, pp. 3069–3074. [Online]. Available: https://www.academia.edu/85939430/The_emoti_chair

[9] S. Merchel and M. E. Altinsoy, “Auditory-Tactile Experience of Music,” in *Music. Haptics*, S. Papetti and C. Saitis, Eds. Springer, Cham, 2018, pp. 123–148. [Online]. Available: http://link.springer.com/10.1007/978-3-319-58316-7_7

[10] S. C. Aker, H. Innes-Brown, K. F. Faulkner, M. Vatti, and J. Marozeau, “Effect of audio-tactile congruence on vibrotactile music enhancement,” *J. Acoust. Soc. Am.*, vol. 152, no. 6, pp. 3396–3409, dec 2022. [Online]. Available: https://asa.scitation.org/doi/10.1121/10.0016444

[11] F. Fontana, F. Avanzini, H. Järveläinen, S. Papetti, G. Klauer, and L. Malavolta, “Rendering and Subjective Evaluation of Real vs. Synthetic Vibrotactile Cues on a Digital Piano Keyboard,” in *Proc. Sound Music Comput. Conf.*, Maynooth, Ireland, jan 2015. [Online]. Available: https://doi.org/10.5281/zenodo.3694948{#}.Y9JyubOYJXg.mendeley

[12] C. Saitis, H. Järveläinen, and C. Fritz, *The Role of Haptic Cues in Musical Instrument Quality Perception*. Cham: Springer International Publishing, 2018, pp. 73–93. [Online]. Available: https://doi.org/10.1007/978-3-319-58316-7_5

[13] G. W. Young, D. Murphy, and J. Weeter, *A Functional Analysis of Haptic Feedback in Digital Musical Instrument Interactions*. Cham: Springer International Publishing, 2018, pp. 95–122. [Online]. Available: https://doi.org/10.1007/978-3-319-58316-7_6

[14] S. Papetti, H. Järveläinen, and S. Schiesser, “Interactive vibrotactile feedback enhances the perceived quality of a surface for musical expression and the playing experience,” *IEEE Transactions on Haptics*, vol. 14, no. 3, pp. 635–645, 2021.

[15] J. A. Russell, “A circumplex model of affect.” *J. Pers. Soc. Psychol.*, vol. 39, no. 6, pp. 1161–1178, 1980.

[16] M. Zentner and T. Eerola, “Self-report measures and models,” in *Handbook of Music and Emotion: Theory, Research, Applications*, P. Juslin and J. Sloboda, Eds. Oxford University Press, 2010.

[17] R. T. Dean, F. Bailes, and E. Schubert, “Acoustic intensity causes perceived changes in arousal levels in music: an experimental investigation.” *PLoS One*, vol. 6, no. 4, p. e18591, apr 2011.

[18] B. W. Vines, C. L. Krumhansl, M. M. Wanderley, and D. J. Levitin, “Cross-modal interactions in the perception of musical performance.” *Cognition*, vol. 101, no. 1, pp. 80–113, 2006.

[19] C. Chapados and D. J. Levitin, “Cross-modal interactions in the experience of musical performances: Physiological correlates,” *Cognition*, vol. 108, no. 3, pp. 639–651, sep 2008.

[20] H. Järveläinen, “Audiovisual perception of arousal, valence, and effort in contemporary cello performance,” in *Proceedings of the Sound and Music Computing Conference*, Malaga, Spain, 2019, pp. 511 – 518.

[21] ——, “Sound-Accompanying Movements Enhance Perceived Arousal in Music Performance Even from a Distance,” in *Proceedings of the Sound and Music Computing Conference*, 2022. [Online]. Available: https://doi.org/10.5281/zenodo.6797851

[22] H. Järveläinen and E. Larrieux, “Dataset: Vibrotactile feedback enhances perceived arousal and listening experience in music,” 2023. [Online]. Available: https://doi.org/10.5281/zenodo.7802367

[23] J. O. Ramsay and B. W. Silverman, *Functional Data Analysis*, ser. Springer Series in Statistics. New York, NY: Springer New York, 2005.

[24] R Core Team, *R: A Language and Environment for Statistical Computing*, R Foundation for Statistical Computing, Vienna, Austria, 2021. [Online]. Available: https://www.R-project.org/

[25] H. Rue, S. Martino, and N. Chopin, “Approximate Bayesian inference for latent Gaussian models by using integrated nested Laplace approximations,” *J. R. Stat. Soc. Ser. B (Statistical Methodol.*, vol. 71, no. 2, pp. 319–392, apr 2009.

[26] “Functional ANOVA using INLA,” https://www.r-bloggers.com/2012/02/functional-anova-using-inla-update/, note = .

[27] R. T. Dean and F. Bailes, “Time series analysis as a method to examine acoustical influences on real-time perception of music,” *Empir. Musicol. Rev.*, no. 4, pp. 152–175.

[28] X. Fu and L. Riecke, “Effects of continuous tactile stimulation on auditory-evoked cortical responses depend on the audio-tactile phase,” 2022 (a preprint not certified by peer-review). [Online]. Available: https://doi.org/10.1101/2022.12.05.519195

[29] Y. De Pra, S. Papetti, F. Fontana, and E. Tiberi, “An Open-Source Robotic Tool for the Simulation of Quasi-Static Finger Pressing on Stationary and Vibrating Surfaces,” *IEEE transactions on haptics*, vol. 14, p. 273–278, 2021.

[30] R. T. Verrillo, “Vibrotactile thresholds measured at the finger,” *Perception & Psychophysics*, vol. 9, no. 4, pp. 329–330, 1971.