Auditory and Tactile Perception of Musical Intervals: A Pilot Study

Davide Deiana
Next Lab
Universita` Campus Bio-Medico di Roma
Rome, Italy
d.deiana@unicampus.it

Alessandro Ansani
...
University of Jyväskylä
Jyväskylä, Finland
alessandro.a.ansani@jyu.fi

Giovanni Di Pino
Next Lab
Universita` Campus Bio-Medico di Roma
Rome, Italy
g.dipino@unicampus.it

Domenico Formica
Neurorobotics Lab
Newcastle University
Newcastle Upon Tyne, UK
Domenico.Formica@newcastle.ac.uk

Charles Spence
Crossmodal Research Lab
University of Oxford
Oxford, UK
charles.spence@psy.ox.ac.uk

Nicola Di Stefano
Institute of Cognitive Sciences and Technologies
National Research Council of Italy
Rome, Italy
nicola.distefano@istc.cnr.it

***Abstract*—The perception of consonance and dissonance has been widely explored in the auditory domain, but little is known about how consonant/dissonant frequency ratios are processed in other sensory modalities, such as touch. In this study, we examined whether consonance-like properties extend to touch by comparing participants' ratings of pleasantness and fusion for both vibrotactile and auditory stimuli. The two sensory modalities were assessed in separate experimental sessions, with auditory stimuli presented dichotically and vibrotactile stimuli delivered as sinusoidal displacements to the participants' hands. Our results indicate that participants' ratings of pleasantness remained consistent across modalities, whereas fusion ratings followed different patterns in audition and touch, influenced by stimulus frequency. Additionally, in the auditory domain, pleasantness and fusion ratings were correlated while, in the tactile domain, pleasantness and fusion exhibited a U-shaped relationship and a relatively low correlation. Our results reveal consistencies in perception across audition and touch, highlighting the need for further investigation into how vibration pairs are processed and perceived across the senses. This research paves the way for future studies on cross-modal consonance.**

***Index Terms*—component, formatting, style, styling, insert.**

I. Introduction

The perception of consonance and dissonance has long been a central topic in music theory and cognition, with origins tracing back to ancient Greek theories of harmonic relationships. Consonant intervals, characterized by simple integer frequency ratios (e.g., octave 2:1, perfect fifth 3:2), are generally perceived as more stable and pleasant, whereas dissonant intervals, with more complex ratios, tend to evoke tension or instability (-[1] for a review).

Theories of consonance and dissonance have been grounded in both physical and perceptual principles. The Pythagorean tradition emphasized the mathematical simplicity of consonant ratios. For instance, Euler (1739) proposed the gradus suavi- tatis [2], a mathematical measure to quantify the pleasantness of musical intervals based on their harmonic simplicity. Later models incorporated psychoacoustic explanations, such as Helmholtz's (1954, [3]) theory of roughness and Terhardt's (1984, [4]) harmonicity framework. More recent research in cognitive science has linked consonance to the efficiency of neural encoding ~~efficiency~~ and predictive processing, suggesting that the brain may favor simpler frequency relationships due to their biolog-ical and computational advantages [5].

Despite extensive research on consonance and dissonance in the auditory domain, relatively little is known about how simple versus complex integer ratios are perceived in other sensory modalities, with no theorization efforts being made in this direction. The few existing studies have examined the perception of auditory frequencies in the tactile domain [6], [7], while others have focused on sensory substitution or augmentation devices that translate auditory information into tactile stimuli (e.g., [8]). Yoo et al. [6] investigated the perception of consonance in superimposed frequencies presented both auditorily and haptically. Their findings indicate that participants could reliably evaluate the degree of consonance in vibrotactile chords. Moreover, for each given frequency, consonance ratings were lowest when the frequency ratio was slightly greater than 1, after which they increased monotonically—aligning with known auditory perception patterns of pure tones (e.g., [9]).

In this pilot study, we ~~aim to~~ contribute to the largely underexplored topic of the tactile perception of consonant and dissonant ratios. In contrast to Yoo et al. [6], our study used auditory/tactile stimuli generated by presenting two different frequencies separately to distinct channels (e.g., ears and hands). For example, a perfect fifth was created by presenting two frequencies in a 3:2 ratio (e.g., 90 Hz and 60 Hz) separately to either the ears or the hands. We investigate whether participants' ratings of pleasantness and fusion for consonant and dissonant ratios are consistent within and across modalities. Additionally, we examine whether non-auditory perception of interval ratios aligns with known auditory pref-

Identify applicable funding agency here. If none, delete this.

erences. By addressing these questions, we seek to provide preliminary insights into the potential multisensory nature of consonance and dissonance perception, paving the way for future research within audition and beyond.

II. MATERIALS AND METHODS

A. Participants

Ten participants (mean age 32.2 ± 7.5 years) were enrolled in this experiment, and all provided written informed consent priorbef to theire participation. The study procedures adhered to the Declaration of Helsinki and subsequent amendments and were approved by the Ethical Committee of Universita` Campus Bio- Medico di Roma (HUROB protocol).

B. Experimental protocol

The experimental protocol aimed at gathering pleasantness and fusion ratings for pairs of sinusoidal pairs that were presented either as auditory or tactile stimuli. Both sensory modalities are sensitive to physical mechanical pressure in the form of oscillations, with a broad bandwidth: from 20 to 20k Hz for sound and from 0.3 to 1000 Hz [10] for touch. Therefore, we decided to use stimuli that rangeding between 40 and 160 Hz.

Each sensory modality was evaluated in a dedicated experimental session. The two sessions were performed on the same day and were separated by a short pause to minimize fatigue. The order of the two blocks was counterbalanced so that half of the participants began with the tactile session while the other half started with the auditory session.

Each session followed the same structure (see Figure 1). Par—ticipants sat in a comfortable chair and received instructions about the task. Prior to the session, the protocol was explained, along with the items participants were invited to rate, namely pleasantness and fusion. Pleasantness is one of the most frequently used constructs in studies investigating consonance perception [1]. Fusion – originally introduced by Stumpf [11],

[12] in reference to auditory percepts - is typically conceived of as the sensory correlate of the merging of multiple tones into a unique percept. Notably, "perceived as one", is one of the criteria participants of the study by Yoo et al. (2014) freely chose to best characterize consonant intervals.

Three random tactile and auditory stimuli were delivered before the start of the experimental task to allow participants to familiarize themselves with the ratings and the stimuli. Thereafter, in the main task, the participants had to evaluate 72 trials. Each trial consisted of a 3-second-long stimulation, either of a vibrotactile or auditory interval, followed by the rating task in which the participants had to judge pleasantness and fusion of the stimulus using a Visual Analog Scale ranging from 1 to 100.

In the tactile condition, the stimulation consisted of skin deformation using two high-performance linear actuators (Mini-shaker Type 4810, BRÜEL & KJÆR). Each motor was placed
on the distal part of the index and middle fingers of each hand (Figure 2). The two motors were activated simultaneously but produced sinusoidal displacements at different frequencies.

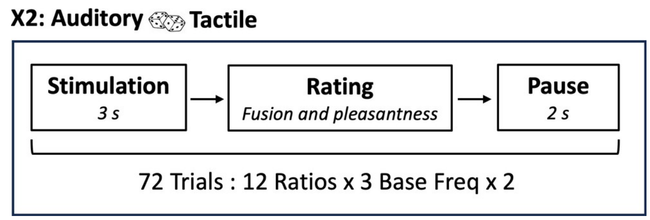

*Fig. 1: Schematic of the experimental protocol. The two sensory modalities (i.e., aAuditory and tTactile) were investigated in two separate experimental sessions, in counterbalanced order. Each session consisted of 72 sub-trials, defined as a
3 s stimulation, a rating phase, and a 2 s pause. Twelve ratios (Table I) were evaluated, generated from three different base frequencies. This process led to 36 different pairs of vibrations. Each pairs of vibrations was repeated twice to avoid a laterality effect (e.g., 40 Hz applied to the right, 80 Hz to the left, and vice versa).*

Insulating headphones were worn by participants in order to minimize any sound feedback from the actuators.

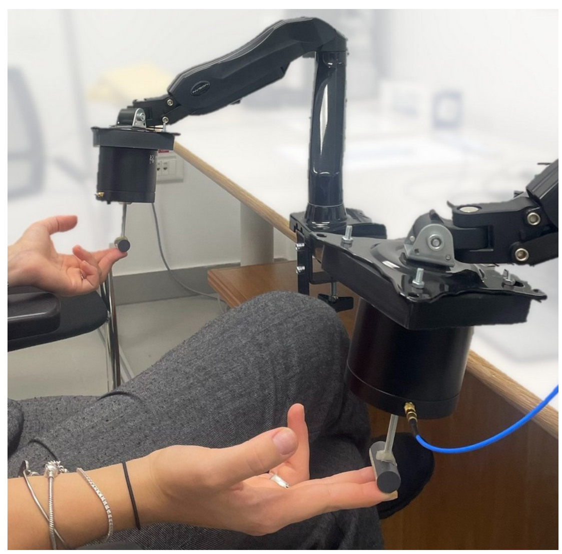

Fig. 2: Experimental setup of the tactile condition. Two linear actuators (Mini-shaker Type 4810, BRÜEL & KJÆR) were positioned in contact with the distal part of the index and middle fingers, each on a different hand. During each stimulation, both actuators were activated simultaneously, following a sinusoidal-like pattern with two different frequencies characterized by a precise ratio (Table I).

In the auditory condition, dichotic pure tones were delivered through commercial headphones (Sony WH-1000XM4). The sound intensity was personalized at the beginning of the session by each participant, who was instructed to set the intensity at a clearly perceptible and comfortable level.

Each stimulation consisted of a sinusoidal pair characterized by specific frequency ratios. A total of 12 ratios were evaluated (see Table I). These ratios are well-known in the context of auditory perception, representing all of the possible intervals within an octave. Three base frequencies (40, 60, and 80

Hz) were selected, resulting in 36 unique stimulus pairs (i.e., 12 ratios × 3 base frequencies). Each frequency pair was presented twice, alternating the side of presentation of each frequency (i.e., the ~~pair~~ 40 Hz - 80 Hz pair was presented both to right and left and left and right). This procedure yielded a total of 72 stimulations per session.

The experimental procedure was managed through a personal computer placed in front of the participant, featuring a custom graphical user interface (GUI) developed in C++ within the Qt Creator 8.0.1 environment. Participants rated pleasantness and fusion using two distinct sliding bars. At the ~~beginning~~ start of each trial, the cursor on each sliding bar was initially positioned at the neutral (central) point. The ~~P~~participants were instructed to move the cursor to the right if the stimula~~-~~tion was perceived as pleasant or fused, and to the left if it was unpleasant or not fused. The magnitude of displacement from the central position reflected the strength of their perception. Stimuli were presented in a randomized order.

Ratio	Base Frequency (Hz)		
	40 Hz	60 Hz	80 Hz
1/2	80.00	120.00	160.00
15/16	42.67	64.00	85.33
8/9	45.00	67.50	90.00
5/6	48.00	72.00	96.00
4/5	50.00	75.00	100.00
3/4	53.33	80.00	106.67
32/45	56.25	84.38	112.50
2/3	60.00	90.00	120.00
5/8	64.00	96.00	128.00
3/5	66.67	100.00	133.33
9/16	71.11	106.67	142.22
8/15	75.00	112.50	150.00

TABLE I: The 12 frequency ratios used in the study. For each base frequency (40, 60, and 80 Hz), the corresponding frequency that results from the given ratio is shown. All values are in Hz.

III. Data Analysis

First, in order to analyze the relationships of our dependent variables (i.e., pleasantness and fusion) within and across modalities, we computed the distance correlation indices (i.e., dCor) via the energy package in R [13]. This metric was preferred over Spearman's ρ as it does not assume monotonicity. The index ranges between 0-1, 0 indicating complete independence [14], [15]. As a second step, we fitted two Linear Mixed Models (LMMs) including both linear and quadratic terms to capture potential non-linear relationships between the ratio and our dependent variables (i.e., pleasantness and fusion). Two- and three-way interactions were specified to examine whether the effect of the ratio varied across modalities and base frequencies. Participants were modeled as random intercepts to account for repeated measures. The models were estimated via the lme4 package in R [16]. The formula in Wilkinson's notation was:

$$DV \sim (\text{ratio}+\text{ratio}^2)\cdot\text{modality}\cdot\text{base frequency}+(1 \mid \text{Participant}) \quad (1)$$

IV. Results

The distance correlation of the pleasantness and fusion scores was $dCor = 0.38$, $p < .001$ in the auditory condition and $dCor = 0.18$, $p < .001$ in the tactile condition. Pleasant- ness was strongly correlated across modalities, $dCor = 0.36$, $p < .001$, whereas fusion was less correlated $dCor = 0.15$, $p < .001$.

When inspecting the estimates of the pleasantness model, ~~we found~~ a significant negative effect of the ratio, $b = -0.63$, $SE = 0.27$, $p = .022$, was found. Notably, ~~we found that suc~~th~~ a~~is decreas-ing trend did not vary across modalities, as signaled by the lack of significance of the interaction term ($b = 0.33$, $SE = 0.39$, $p = .399$). ~~We also found~~ A~~a~~ significant interaction between the ratio and the base frequency ($b = -0.57$, $SE = 0.27$, $p = .039$) was also found. ~~All~~ None of the quadratic terms nor an~~d~~y of their interactions were ~~not~~ significant.

In the fusion model, both the linear decrease ($b = -1.60$, $SE = 0.35$, $p < .001$) and the quadratic term ($b = 0.36$, $SE = 0.11$, $p = .002$) of the ratio effect were significant. Dif- ferently from the pleasantness model, the linear decrease did not remain consistent across modalities ($b = 1.19$, $SE = 0.50$, $p = .018$). However, we found evidence that the relationship between ratio and fusion changed depending on the base frequency. Both the linear ($b = -0.72$, $SE = 0.35$, $p = .042$) and quadratic interaction terms ($b = 0.22$, $SE = 0.11$, $p = .058$) pointed in this direction. Notably, the three-way interaction ratio×modality×base frequency reached significance too ($b = 1.09$, $SE = 0.50$, $p = .030$), thus indicating that the relationship between ratio and modality changed at different base frequency levels. Lastly, the relationship between fusion scores in the two modalities was not consistent across base frequency levels ($b = 0.77$, $SE = 0.37$, $p = .038$). The relationships between ratio, modalities, and base frequency are represented in Figure 3.

V. Discussion

The question of what makes a combination of tones be perceived as consonant or dissonant has fascinated humanity for thousands of years, with some of the earliest theories tracing back to the ancient Greeks. In these early theories, the ratio between the frequencies of two pure tones was considered the basic principle for classifying sounds as consonant or dissonant. Over time, more complex theories have been proposed, and the scope of the study has expanded to include fields such as physiology, cognitive science, and neuroscience.

Despite significant progress and increasingly sophisticated models, simple integer ratios still provide a good approximation and prediction of sound consonance [17]. However, while extensive research has demonstrated the importance of frequency ratios in auditory perception, much less is known about their role in other sensory modalities.

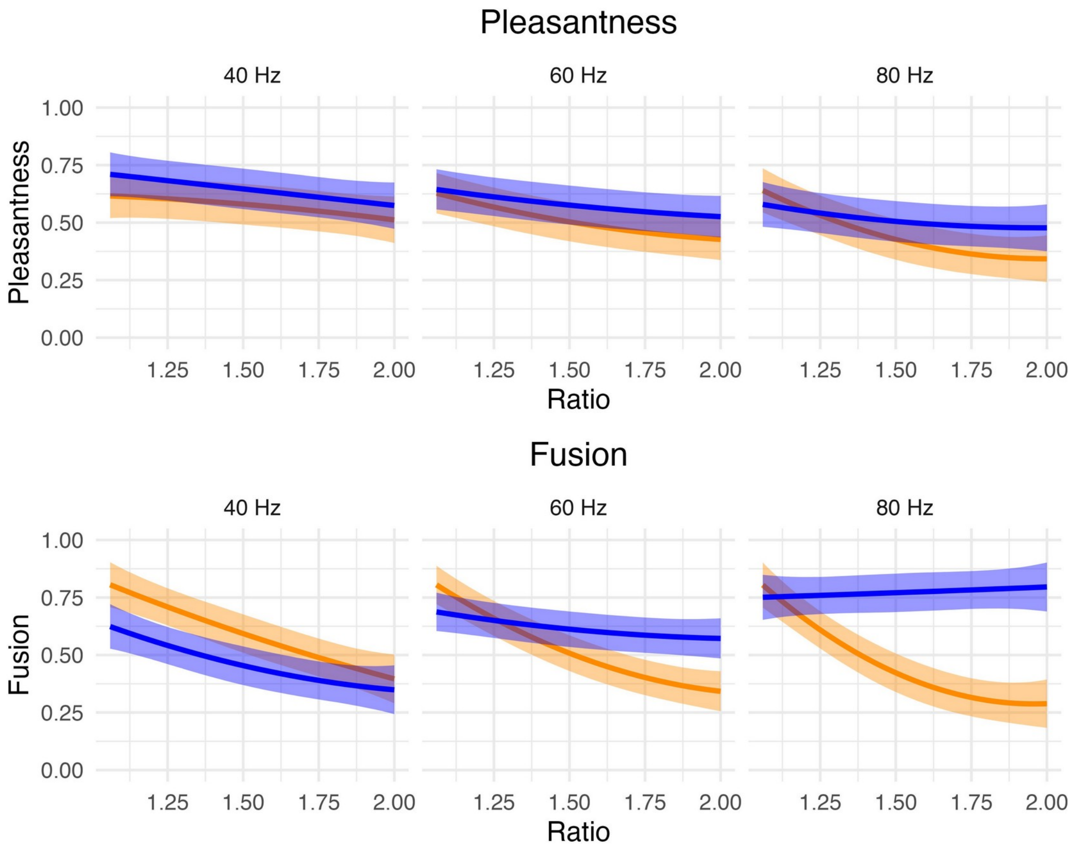

Fig. 3: Results of the Linear Mixed Models. The auditory condition is represented in orange, the tactile condition in blue. The shaded area represents the 95% Confidence Interval.

In this study, ten participants ~~were asked to~~ rated the pleasantness and fusion of vibration pairs presented through both the auditory and tactile modalities. The results reveal~~show~~ed that participants' ratings of pleasantness remained consistent across modalities, whereas their fusion ratings followed different patterns in audition and touch, influenced by stimulus frequency. Additionally, in the auditory domain, pleasantness and fusion ratings were correlated, aligning with previous research on consonance perception (e.g., [18], [19]). In contrast, in the tactile domain, pleasantness and fusion exhibited a U-shaped relationship and a relatively low correlation.

The ratings of auditory intervals obtained in our study do not align with those reported in similar studies (e.g., [9]). These discrepancies may ~~be~~ only be partially attributed to the narrow and relatively low-frequency range used in our protocol and the limited number of participants. One additional explanation might be related to perceived intensity which, within our frequency range, is influenced by frequency. The subjective equality curves show a steep slope, where even a 40 Hz difference may require a 10–20 dB adjustment for two pure tones to be perceived at the same intensity [20]. This effect may have impacted participants' judgments, particularly at higher frequency ratios. Since we maintained a constant sound pressure level (SPL) for both frequencies, one tone may have been perceived as more dominant, potentially influencing how participants evaluated the stimulus as a combination of tones.

Regarding touch, previous research has indicated that pleasantness ratings increase significantly beyond 130–150 Hz [6], whereas most of our stimuli fell within the 40–120 Hz range. Notably, our study methodology differs substantially from [6], where participants experienced the two vibrations superimposed. In contrast, in our study, each vibrotactile frequency was presented to a different hand. These differences may partially explain the divergent results. Nonetheless, based on Yoo's findings [6], future studies should include higher-frequency stimuli, ideally up to 250–300 Hz, to explore the combination of intervals extending beyond the octave.

Finally, the influence of frequency on ratings in the tactile domain may stem from the differing effects of vibration pairs at lower versus higher frequencies. The transition from 40 to 80 Hz can significantly impact the sensation, shifting from distinguishable pulse trains to indistinct vibratory pulses.

In conclusion, our pilot study provides novel insights into consonance from a multisensory perspective. The correlation of pleasantness ratings across senses is particularly noteworthy, as participants are likely familiar with auditory stimuli but may be encountering "tactile harmonies" for the first time. Examining touch and audition through vibration pairs offers a promising approach to understanding shared mechanisms in sensory processing across modalities.

REFERENCES

[1] N. Di Stefano, P. Vuust, and E. Brattico, “Consonance and dissonance perception. A~~a~~ critical review of the historical sources, multidisciplinary findings, and main hypotheses,” *Physics of Life Reviews*, vol. 43, pp. 273–304, 2022.
[2] L. Euler, *Tentamen novae theoriae musicae []*. Ex Typographia Academiae Scientiarum, 1739.
[3] H. L. Helmholtz, “On the sensations of tone as a physiological basis for the theory of music,” 1954.
[4] E. Terhardt, “The concept of musical consonance: A link between music and psychoacoustics,” *Music P~~p~~erception*, vol. 1, no. 3, pp. 276–295, 1984.
[5] A. Tabas, M. Andermann, V. Schuberth, H. Riedel, E. Balaguer- Ballester, and A. Rupp, “Modeling and ~~meg~~MEG evidence of early con- sonance processing in auditory cortex,” *PLoS C~~c~~omputational B~~b~~iology*, vol. 15, no. 2, p. e1006820, 2019.
[6] Y. Yoo, I. Hwang, and S. Choi, “Consonance of vibrotactile chords,” *IEEE Transactions on Haptics*, vol. 7, no. 1, pp. 3–13, 2014.
[7] R. Okazaki, T. Hachisu, M. Sato, S. Fukushima, V. Hayward, and H. Kajimoto, “Judged consonance of tactile and auditory frequencies,” in *2013 World Haptics Conference (WHC)*. IEEE, 2013, pp. 663–666.
[8] M. Karam, C. Branje, G. Nespoli, N. Thompson, F. A. Russo, and D. I. Fels, “The emoti-chair: A~~a~~n interactive tactile music exhibit,” in *CHI’10 Extended Abstracts on Human Factors in Computing Systems*, 2010, pp. 3069–3074.
[9] A. Kameoka and M. Kuriyagawa, “Consonance theory part i: Conso- nance of dyads,” *The Journal of the Acoustical Society of America*, vol. 45, no. 6, pp. 1451–1459, 1969.
[10] G. W. Young, D. Murphy, and J. Weeter, “Haptics in music: T~~t~~he effects of vibrotactile stimulus in low frequency auditory difference detection tasks,” *IEEE T~~t~~ransactions on H~~h~~aptics*, vol. 10, no. 1, pp. 135–139, 2016.
[11] C. Stumpf, “Tonpsychologie, vol. ii,” *Leipzig: Hirzel*, 1890.
[12] ——, *Beiträge zur akustik und musikwissenschaft*. JA Barth, 1909, vol. 4.
[13] M. Rizzo, G. Székely *et al.*, “E-statistics: Multivariate inference via the energy of data,” *energy*, vol. 1, p. 11, 2022.
[14] G. J. Székely, M. L. Rizzo, and N. K. Bakirov, “Measuring and testing dependence by correlation of distances,” 2007.
[15] G. J. Székely and M. L. Rizzo, “Brownian distance covariance,” 2009.
[16] D. Bates, M. Mächler, B. Bolker, and S. Walker, “Fitting linear mixed-effects models using lme4,” *Journal of statistical software*, vol. 67, pp. 1–48, 2015.
[17] N. Di Stefano, “Cultural accounts of consonance perception. ~~a~~A Lakatos Ian approach to save Pythagoras,” *Music Perception: An Interdisciplinary Journal*, vol. 42, no. 2, pp. 165–176, 2024.
[18] M. Guernsey, “The role of consonance and dissonance in music,” *The American Journal of Psychology*, pp. 173–204, 1928.
[19] L. A. DeWitt and R. G. Crowder, “Tonal fusion of consonant musical intervals: The oomph in Stumpf,” *Perception & Psychophysics*, vol. 41, no. 1, pp. 73–84, 1987.
[20] S. Merchel and M. E. Altinsoy, “Psychophysical comparison of the auditory and tactile perception: a survey,” *Journal on Multimodal User Interfaces*, vol. 14, no. 3, pp. 271–283, 2020.