Interactive Vibrotactile Feedback Enhances the Perceived Quality of a Surface for Musical Expression and the Playing Experience

Stefano Papetti, *Member, IEEE*, Hanna Järveläinen, and Sébastien Schiesser

***Abstract*—An advanced multi-touch sensor surface aimed at musical expression was recently equipped by the authors with interactive multi-point localized vibrotactile feedback. Using such interface, a subjective assessment was conducted that measured how the presence and type of vibration affect the perceived quality of the device and various attributes related to the playing experience. Two clearly distinct sound settings each with three vibrotactile feedback strategies were tested. At each trial, the task was to play freely while comparing two related setups which used the same sound setting and differed only in the presence/absence of vibration. Independent of the sound setting, as compared to the respective non-vibrating setups, vibrations conveying frequency and amplitude dynamics cues coherent with the player's gesture and/or sonic feedback had the most positive effect. Vibrotactile feedback especially improved the enjoyment of playing and the perceived potential for musical expressivity.**

***Index Terms*—Digital musical interface, haptic surface, music performance, quality perception, vibrotactile feedback.**

I. INTRODUCTION

Recent efforts in human-centered and technological research aimed to gain a better understanding of how vibrotactile signals can be conveyed to the user [1], [2], as well as to develop novel interactive surfaces yielding a rich haptic experience. On touchscreens, force sensing is being investigated as an important feature for expanding the palette of input gestures that can be detected [3]. With regard to haptic cues provided by interactive surfaces, localized feedback and friction modulation are now regarded as key missing links to a more natural haptic response [4], [5].

In parallel, over the last few years interactive surfaces have become popular interfaces for musical interaction. Notable early examples are the successful ReacTable [6] and Lemur[1] devices. This process has been assisted by the diffusion of touchscreen technology in mobile devices such as smartphones and tablets – with their increasing computing power and multimedia capabilities – and by the growing multitude of high quality musical applications developed specifically for those platforms (e.g., AniMoog[2] and the recent software version of the Lemur[3]). A major advantage of touchscreen-based vs hardware user interfaces is that the former offer a space that can be freely configured by the developer and possibly by the user (as in the Lemur and TouchOSC[4] apps), opening to unexplored possibilities in various musical interactions (e.g., music performance, music recording/mixing, and control of musical processes). Those interfaces, however, still fall short of establishing a rich physical exchange with the user.

In this perspective, several studies in the emerging field of Musical Haptics [7] aspired on the one hand to design next-generation digital musical interfaces yielding haptic feedback [8], and on the other hand to assess if and how haptic feedback is relevant to the perceived quality of musical instruments, to the performer's experience and performance, and to the resulting musical outcome [9]–[12].

Among traditional musical instruments, the haptic response of the piano and the violin have been mainly studied, generally finding that haptic cues are relevant to their perceived quality, and may be even more important than auditory feedback for their identification [13]–[16]. Several examples of digital musical instruments yielding haptic feedback are found in the literature, some of which were also qualitatively assessed in relation to the offered haptic cues [17]–[19]. However, only a few studies so far addressed the design and evaluation of haptic surfaces for musical expression [20], [21].

In the framework outlined above, we conducted a subjective assessment making use of a force-sensitive multi-touch surface for musical interaction, which we augmented with multi-point localized vibrotactile feedback. Our study investigated how different types of vibrotactile feedback affect various attributes related to the playing experience and the perceived quality of the interface.

Manuscript received September 14, 2020; revised February 12, 2021; accepted February 15, 2021. Date of publication February 19, 2021; date of current version September 16, 2021. This work was supported in part by project HAPTEEV funded by the Swiss National Science Foundation under
Grant 178972. This article was recommended for publication by Associate Editor Prof. Eckehard Steinbach and Editor-in-Chief Dr. Seungmoon Choi upon evaluation of the reviewers' comments. *(Corresponding author: Stefano Papetti.)*

The authors are with the Institute for Computer Music and Sound Technology, Zurich University of the Arts, 8005 Zurich, ZH, Switzerland (e-mail: stefano.papetti@zhdk.ch; hanna.jarvelainen@zhdk.ch; sebastien.schiesser@zhdk.ch).

Digital Object Identifier 10.1109/TOH.2021.3060625

[1] [Online]. Available: https://en.wikipedia.org/wiki/Lemur_Input_Device

[2] [Online]. Available: https://www.moogmusic.com/products/animoog

[3] [Online]. Available: https://liine.net/en/products/lemur/

[4] [Online]. Available: https://hexler.net/products/touchosc

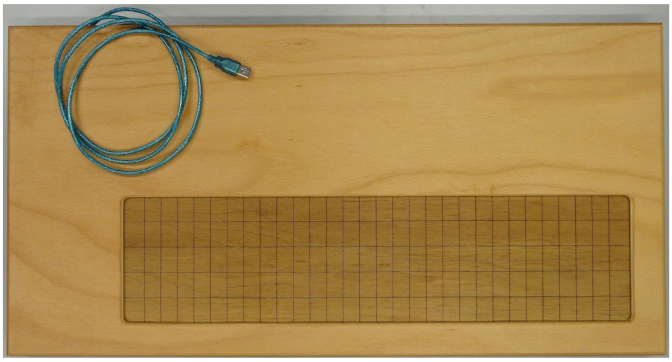


Fig. 1. The HSoundplane, the haptic musical interface used in the experiment.

II. VIBROTACTILE AND SOUND FEEDBACK SYSTEM

The HSoundplane, shown in Fig. 1, is a haptic musical interface prototype yielding multi-point, localized vibrotactile feedback across its surface. The device was used in the experiment to let participants experience various types of sound and vibration feedback while playing. It is based on the Madrona Labs Soundplane [22]: an advanced musical controller offering a large ($55.5 \times 14\,\mathrm{cm}$) multi-touch and pressure-sensitive surface.Its patented audio-rate capacitive sensing technology results in tracking times in the order of a few ms — a fraction of the lag offered by the fastest touchscreen technology currently available [23].

The original device was augmented with interactive vibrotactile feedback: an actuator layer based on piezoelectric discs and a software synthesis system were developed, resulting in the HSoundplane prototype (where ‘H’ stands for ‘haptic’). Its hardware design has been described in detail in [24], therefore only essential information is reported here.

In order to drive the piezo discs with standard audio signals, custom amplifying and routing electronics were designed, based on the following components: Texas Instruments DRV2667 piezo drivers, serial-to-parallel shift registers with output latches of the 74HC595 family, high voltage MOSFET relays, four slave microcontrollers and one master microcontroller. The piezo elements were arranged on a flexible PCB foil in a 30×5 matrix configuration, matching the tiled pads on the Soundplane’s surface. Since each piezo driver feeds 5 actuators in parallel (one driver per column), particular attention was paid to current consumption and heat dissipation: Eventually, Murata Electronics 7BB-20-6 piezo actuators[5] (resonating at $6.3\,\mathrm{kHz}$) were selected mainly for their small capacitance ($10\,\mathrm{nF}$) while offering a frequency response matching the tactile range (see II-C). The flexible PCB foil connecting the piezo elements was laid on top of a thin rubber sheet with holes corresponding to each piezo element, thus ensuring enough free space to allow for optimal mechanical deflection.

A client software for Mac computers comes with the original Soundplane, which receives multi-touch data sensed by the interface and routes them to other applications using the Open Sound Control (OSC) protocol.[6] Such messages carry absolute x, y coordinates (for position) and pressing force values along the z-axis for each contacting finger. An additional software application was developed making use of Cycling ’74 Max,[7] which receives OSC touch data from the Soundplane application and use them to drive two DSP engines that synthesize and route audio and vibration signals, respectively. The sound and vibration feedback settings designed for the experimental setup are described below.

A. Sound Feedback

Sound was provided to the participants by means of closed-back headphones (Beyerdynamic DT 770 Pro).

The pitch of the audio feedback is controlled along the x-axis according to a chromatic subdivision which maps each pad of the tiled surface to a semitone. Pitch ranges from A2 ($f_0 = 110\,\mathrm{Hz}$) to D5 ($f_0 = 587.33\,\mathrm{Hz}$). Similar to the customary string coloring on the harp, the columns corresponding to C and F tones were painted respectively in red and blue (see Fig. 6), thus providing a clear pitch reference to the participants.

Two types of sonic feedback were designed:

- **Sound 1:** A markedly expressive setting, which responds to subtleties and nuances in the performer’s gesture. It consists in a sawtooth wave filtered by a resonant low-pass and modulated by a vibrato effect (i.e., amplitude and pitch modulation).
  y-axis control: The vibrato intensity varies exponentially along the y-axis from no-vibrato (bottom) to strong vibrato (top).
  z-axis control: The filter cutoff frequency is controlled by the applied pressing force (i.e., higher force maps to brighter sound), and so is the sound level (i.e., higher force maps to louder sound).
- **Sound 2:** A setting offering a rather limited sonic palette and no amplitude dynamics. It consists in a simple sine wave to which noise is added depending on the location on the y-axis.
  y-axis control: Moving upwards adds white noise of increasing amplitude, filtered by a resonant band-pass. The filter’s center frequency follows the pitch of the respective tone.
  z-axis control: Pressing force data are ignored, resulting in fixed loudness.

All sounds are processed by a reverb effect which makes the playing experience more acoustic-like.

Audio examples of the two sound types are made available online,[8] demonstrating C3, C4 and C5 tones modulated along the y- and z-axes.

B. Vibrotactile Feedback

Before being routed to the actuators, vibration signals are filtered in the $10{-}500\,\mathrm{Hz}$ range by a 10th-order band-pass, so

---

[5] [Online]. Available: https://www.murata.com/en-eu/products/productdetail?partno=7BB-20-6

[6] [Online]. Available: http://opensoundcontrol.org/

[7] A visual programming environment for multimedia and physical computing: [Online]. Available: https://cycling74.com

[8] [Online]. Available: https://tinyurl.com/HS-sounds

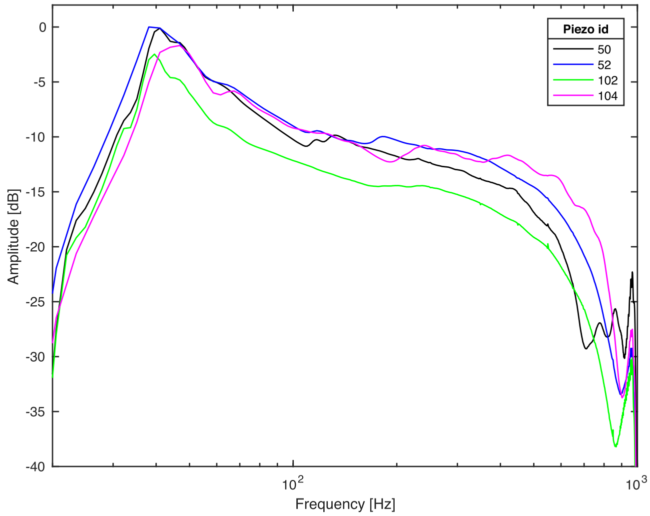


Fig. 2. Normalized vibration frequency response in the $20-1000$ Hz range (FFT size 16384), measured at four exemplary piezo transducers at the 12th (#50, 52) and 22nd column (#102, 104).

as to optimize the actuators' efficiency and consequently the vibratory response of the device, as well as to minimize sound leakage. Any residual sound spillage produced by the actuators was taken care of by the closed-back headphones carrying auditory feedback.

Three vibrotactile strategies were implemented:

- **Sine:** Pure sinusoidal signals are used, whose pitch follows the fundamental of the played tones (f_0 within $110-587.33$ Hz), and whose amplitude is controlled by the intensity of the pressing forces. By focusing vibratory energy at a single frequency component, this setting aims at producing sharp vibrotactile feedback.
- **Audio:** The same sounds generated by the HSoundplane are used to render vibration: the audio signals are also routed to the actuators layer. Vibration signals thus share the same spectrum (within the $10-500$ Hz passband) and dynamics of the related sound. This approach ensures the highest coherence between musical output and tactile feedback, mimicking what occurs on acoustic musical instruments, where the source of vibration coincides with that of sound.
- **Noise:** A white noise signal is used, whose amplitude is fixed. This setting produces vibrotactile feedback that is generally uncorrelated with the auditory one, ignoring any spectral and amplitude cues possibly conveyed by it. The only exception is with Sound 2 and high y-axis values, which results in a similar noisy signal.

The intensity of vibration feedback was set by the authors in a pilot phase, aiming at two main goals: i) sound and vibration intensities had to feel reciprocally consistent; ii) while levels had to be overall comfortable for prolonged use, vibration had to be clearly perceivable even at low force-pressing values [2].

C. Characterization of Vibratory Output

The frequency response of the vibration generation system in the vibrotactile band [25] was measured with a Wilcoxon 736 T accelerometer, stuck with double-side tape at several pads of the HSoundplane's surface, while the underlying piezo transducers were fed with sinusoidal sweeps [26] in the $20-1000$ Hz range. Figure 2 shows the results of measurements performed in correspondence of four exemplary piezo transducers. In general, the frequency responses measured at different locations over the surface are very similar in shape, with a pronounced peak at about 40 Hz. In some cases they show minor amplitude offsets (e.g., the response of piezo #102 in Fig. 2) that can be compensated for at software level.

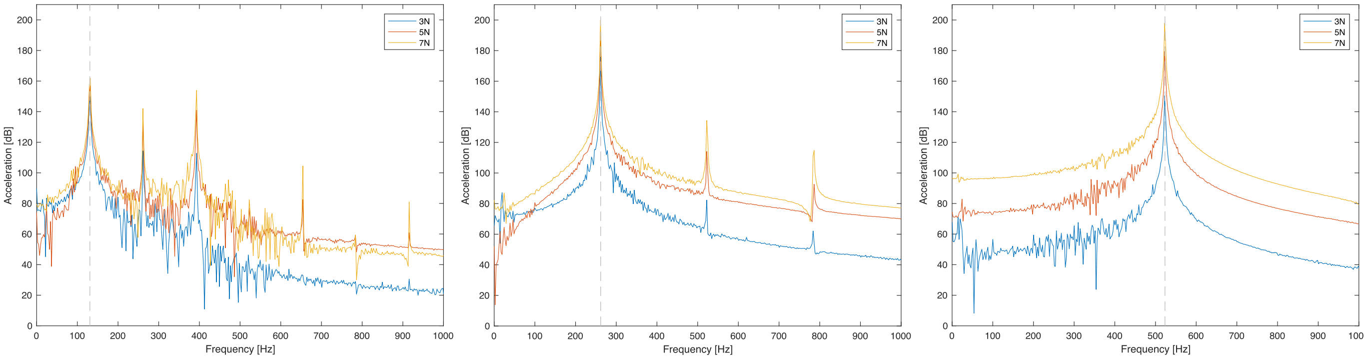
For several reasons, the measured frequency response can only partially render the actual response of the vibrotactile system while in use: First of all, such response describes only the linear regime of the actuators; Also, vibration types Audio and Noise feed the actuators with spectrally complex signals – as opposed to the swept sine signal used for the frequency response measurement – which may give rise to nonlinear distortion; Finally, any load on top of a vibrating object (e.g., a pressing finger) alters its vibration amplitude and spectral energy distribution.

To deal with the above issues, characterization was also performed using the Sine and Noise vibration types used in the experiment, having polar opposites qualities in terms of spectra and dynamics, while measurements were acquired with a self-developed robotic device that simulates a human finger pressing down vertically (up to 20 N) and embeds a PCB 356A17 triaxial accelerometer [27]. The goal was to verify how accurately the original signals feeding the actuators would be rendered while a finger pressed on the HSoundplane's surface, and to measure the generated output power (RMS acceleration).

No characterization was performed with the Audio vibration type, based on the following observations: For Sound 1, the rendered Audio vibration would not offer a clearly defined reference against the original signal, as this is highly variable depending on the applied force (amplitude and filter cutoff modulation) and location along the y-axis (vibrato modulation). Moreover, it is known that tactile waveform discrimination is mainly effective below 100 Hz [28], so sawtooth waves with f_0 in our target range ($110-587.33$ Hz) should be indistinguishable from sine waves (confirmed by our informal tests). Also, in the case of Sound 2, Audio vibration would not provide any further information compared to Sine and Noise characterization.

As in the actual experiment, the test signals were band-pass filtered in the $10-500$ Hz range before feeding the actuators. Vibrations were measured at 10 locations across the HSoundplane's surface (namely, the pads at the intersections of rows #2 and 4 with columns #4, 9, 16, 21, 29), for 3 pressing-force levels (3, 5, 7 N).

The spectral characterization of Sine vibration (Fig. 3) shows that the input signals are generally rendered accurately, however for lower-pitched signals (approximately from F3 downwards, that is for $f_0 \leq 174.61$ Hz) spurious harmonics are generated with amplitude close to that of the fundamental. With regard to vibration intensity, Fig. 4 reports the RMS acceleration of the rendered signals for various tones and pressing forces. Amplitude obviously varies with the pressing force (as designed, see II-B), but also with f_0 — in particular



(a) C3 ($f_0 = 130.81\,\mathrm{Hz}$), row #4 (b) C4 ($f_0 = 261.63\,\mathrm{Hz}$), row #4 (c) C5 ($f_0 = 523.25\,\mathrm{Hz}$), row #4

Fig. 3. Amplitude spectra of rendered Sine vibration with C3, C4, and C5 pitch, for 3, 5, 7 N pressing forces (FFT size 8192 samples; dB re $10^{-6}\,\mathrm{m/s^2}$). The vertical dashed lines mark the frequency of the original signal.

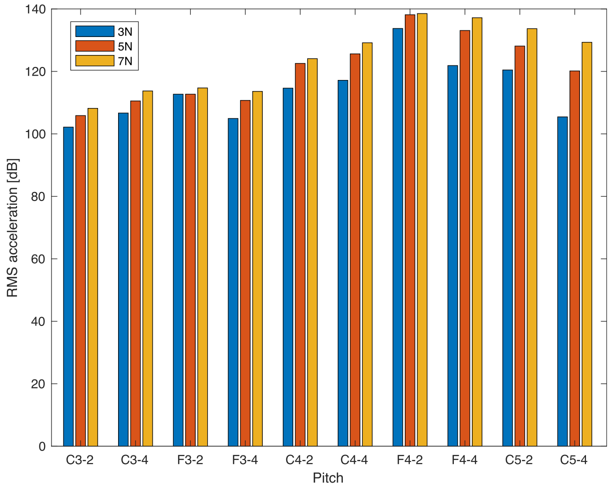


Fig. 4. Vibration amplitude of Sine vibration for different tones measured at rows #2 and #4 (denoted as ‘pitch-2’ and ‘pitch-4’) and 3, 5, 7 N pressing forces (dB re $10^{-6}\,\mathrm{m/s^2}$).

a peak can be observed around F4 ($f_0 = 349.23\,\mathrm{Hz}$). The mean RMS acceleration (dB re $10^{-6}\,\mathrm{m/s^2}$) at the measured locations for 3, 5, 7 N pressing forces is $114\,\mathrm{dB}$ ($\mathrm{sd} = 9.8$), $120.8\,\mathrm{dB}$ ($\mathrm{sd} = 10.7$), $124.2\,\mathrm{dB}$ ($\mathrm{sd} = 11.0$), respectively.

The spectral characterization of Noise vibration (see Fig. 5) is also consistent with the input signal, showing well-preserved and uniformly “colored” wide-band spectra with high-pass cutoff at about 100 Hz, and an amplitude dip just above the cutoff. The main energy peak (spectral centroid) varies slightly with the pressing force: The mean spectral centroids at the measured locations are $307\,\mathrm{Hz}$ ($\mathrm{sd} = 7.6$), $330.6\,\mathrm{Hz}$ ($\mathrm{sd} = 6.8$), and $345.7\,\mathrm{Hz}$ ($\mathrm{sd} = 4.8$) respectively for 3, 5, 7 N pressing forces; overall mean $327.8\,\mathrm{Hz}$ ($\mathrm{sd} = 17.4$). The spectral shapes in Fig. 5 are consistent with the vibration amplitudes at different frequencies shown in Fig. 4. By design, the amplitude of Noise vibration does not change with the applied force, and the HSoundplane accurately reproduces such behavior across its surface and for the tested pressing forces: The overall average RMS acceleration is $125.47\,\mathrm{dB}$ ($\mathrm{sd} = 1.37$).

In general, the intensity of vibration feedback is between $40{-}60\,\mathrm{dB}$ higher than the sensitivity thresholds related to active touch [2]. As a reference, a clear sensation of vibration was reported to arise for stimuli $40\,\mathrm{dB}$ above threshold [29].

III. EXPERIMENT

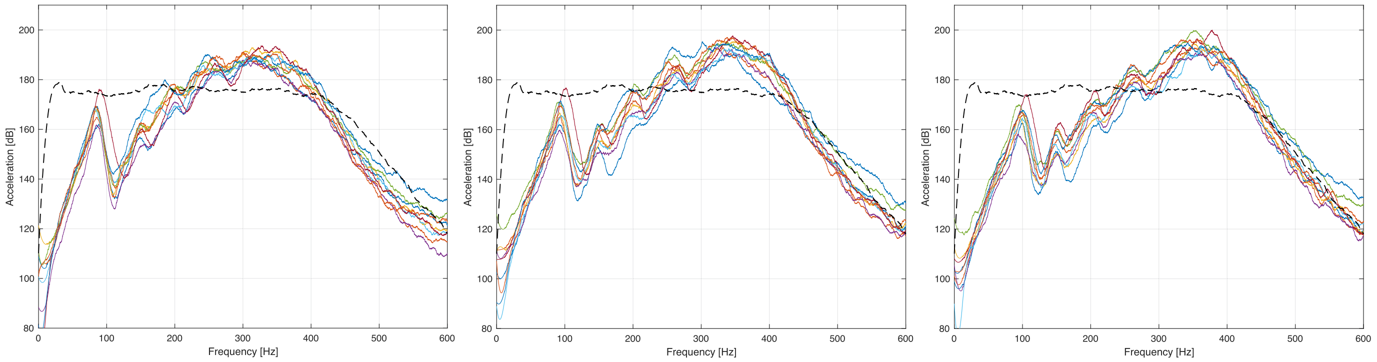
The goal of the experiment was to assess how different types of vibration could affect the perceived quality of the interface and the playing experience. The three available vibration types (see II-B) offered different spectral and dynamics cues resulting in varying degrees of similarity with the audio feedback, thus enabling to determine the importance of the match between sound and vibration. Also, the two available sound settings (see II-A) offered different degrees of variability and expressive potential, allowing to investigate whether the possible effect depends on audio feedback characteristics. The assessments were made by comparing each of the vibrating setups against a respective non-vibrating configuration with the same sound setting.

A. Design, Procedure, and Subjects

The test method was comparison rating with a hidden reference, where the condition under test (vibrating setup) is compared to a reference (non-vibrating setup) as a pseudo-paired comparison [30]. Here the test and reference setups shared the same audio setting, while the test setup additionally provided vibrotactile feedback. The two compared setups were assigned to labels A and B in a balanced way. The three vibration types were crossed with the two sound settings, and ratings were measured on four attributes: *Preference*, *Control and responsiveness* (referred to as *Control*), *Expressive potential* (referred to as *Expression*), and *Enjoyment*.

Participants had to wear closed-back headphones and sit at a table where the HSoundplane, a LCD screen and a mouse were placed (Fig. 6). The LCD screen displayed a mouse-operated GUI with a switch for setup selection (A/B), horizontal slider (s) for assigning ratings, and a timer.

Participants would play the HSoundplane freely for a given time, switching between setups A and B whenever they wished, and then rate the attributes as prompted by the GUI. Ratings for each attribute were given by adjusting a respective slider on a continuous visual analog scale ranging from A



(a) Pressing force 3 N (b) Pressing force 5 N (c) Pressing force 7 N

Fig. 5. Smoothed amplitude spectra of rendered Noise vibration for 3, 5, 7 N pressing forces (FFT size 262144 samples; dB re $10^{-6}\ \mathrm{m/s^2}$). The spectrum of the original signal is represented by black dashed lines, while colored lines report the responses measured at ten locations distributed across the surface (namely, the pads at the intersections of rows #2 and 4 with columns #4, 9, 16, 21, 29).

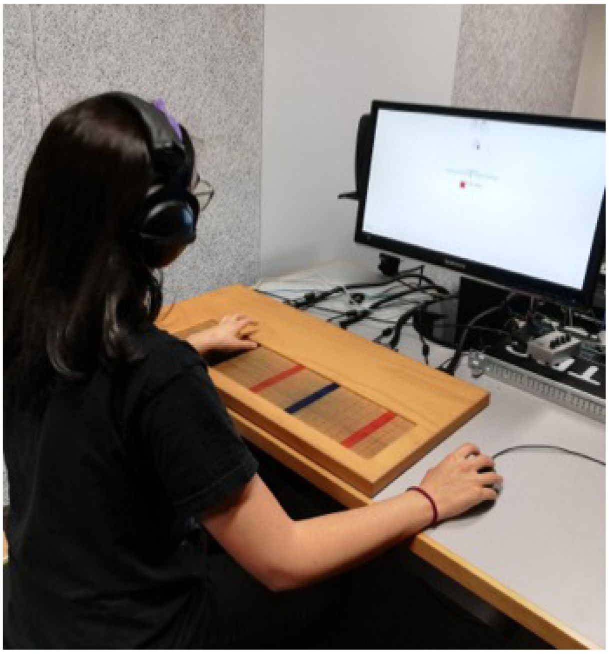


Fig. 6. The experimental setting.

(left) to B (right) to reflect the degree of preference in terms of the given attribute (in case of perceived equality between A and B the slider would be set to the midpoint). The same rating scale was used for all attributes.

All 4 (attributes) × 3 (vibration types) × 2 (sound types) factor combinations were evaluated twice. Trials were grouped in blocks based on sound type, each including a full round of attribute-vibration combinations. The vibrating/non-vibrating setups were assigned to labels A and B also block-wise, and the block presentation order was balanced. Block-wise grouping was meant to avoid continually disrupting the musical experience that would result from frequently changing sound or the assignment of vibration to setups A/B. To summarize, one measurement block was made of 6 playing periods of the same sound type. The entire experiment consisted of four such blocks, two for each sound.

The sliders for rating *Control*, *Expression*, and *Enjoyment* were presented together, while *Preference* was rated separately. The rationale for such organization is that rating all four attributes together after a single playing period might have introduced a cognitive bias (e.g., very similar ratings for all attributes); conversely, rating all attributes separately would have resulted in 48 trials each involving a playing period, making the experiment excessively long.

A two-minute playing period was recommended for the *Preference* trials, and four minutes for the three-attribute trials. The nominal duration of the experiment therefore was 72 minutes, however participants were typically finished within one hour.

Before the experiment, participants were briefed about the procedure and could familiarize for 5-10 minutes with the instrument in the non-vibrating configuration, testing both sound types. In order to avoid bias, they were merely informed that the two setups were different, without further explanation.

All 29 participants – 7 males and 22 females, aged 18-48 years (M = 25.4, sd = 7.1) – were professional musicians or music students. Their main instrument was either a keyboard or a string instrument, on which they had on average 17 years of experience. Roughly one third of the participants had significant experience with electronic musical instruments, mostly synthesizers, or digital musical interfaces.

B. *Results*

The data collected in the experiment (responses and gestural data) as well as the software scripts used to analyze them are made available in an open-access repository.[9]

1) Slider Response Data: The continuous slider scale ratings were mapped to the closed interval [0,1], where 1 indicates maximal preference for the vibrating setup and 0 maximal preference for the non-vibrating setup, and 0.5 the point of perceived equality.

For both sound types, the large majority of participants gave positive ratings to Sine and Audio vibrations: Sine vibrations were rated positively by N = 22 and N = 19 subjects in combination with Sound 1 and Sound 2, respectively; Audio

---

[9] [Online]. Available: https://doi.org/10.5281/zenodo.4028637

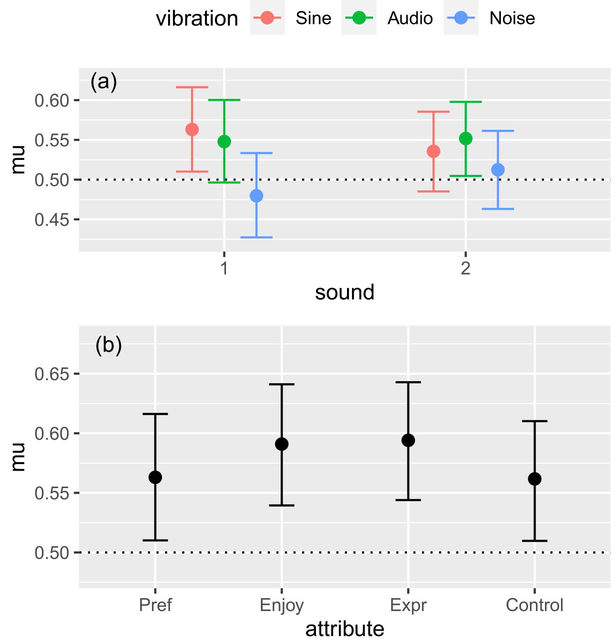


Fig. 7. Marginal effects; estimated μ parameters with 95% Credible Intervals (N=29). (a) Interaction between vibration and sound type; (b) Effect of vibration on the evaluated attributes.

vibrations received positive ratings from N=18 and N=22 participants with Sound 1 and Sound 2. The least preferred feedback was Noise, which was rated positively by N=14 and N=13 participants (i.e., still by nearly half of the subjects) respectively for Sound 1 and Sound 2.

As is typical for visual analog scale data, the dataset showed strong deviations from normality due to the restricted interval, hence analysis of variance could not be applied. The results were therefore analysed using a zero-one-inflated beta (ZOIB) model, whose parameters were estimated with Bayesian methods [31]–[33]. The *brms* package for R was used, which allows to specify models using the same syntax as for general linear models [34].

The ZOIB distribution is well-suited for closed-interval data such as slider ratings, and models the data in two components: a beta distribution models the open interval (0, 1), and a Bernoulli distribution the binary responses $\{0, 1\}$ (i.e., the sliders' endpoints). On the downside, four parameters are needed to describe the distribution: the mean (μ) and precision (ϕ) of the beta distribution, the probability of a binary $\{0, 1\}$ outcome (zoi), and the conditional probability of outcome $\{1\}$ (coi). Each of these parameters may be modeled by a number of main effects and interactions of the manipulated factors. Based on descriptive analysis of the data, we formulated the following model:

$$\begin{aligned}
\text{rating } \mu &\sim \text{sound} * \text{vibration} + \text{attribute} \\
&\quad + (\text{sound} + \text{vibration}|\text{sID}) \\
\phi &\sim \text{vibration} + \text{sound} + \text{attribute} \\
\text{zoi} &\sim \text{vibration} + \text{sound} + \text{attribute} \\
\text{coi} &\sim \text{vibration} + \text{sound} + \text{attribute}
\end{aligned}$$

TABLE I

SUMMARY OF THE ZOIB MODEL FIT. COLUMNS: PARAMETER NAME, ESTIMATE, LOWER AND UPPER BOUNDS OF 95% CI. ESTIMATES ARE DESCRIBED ON THE SCALE OF THEIR RESPECTIVE *LINK FUNCTIONS*, I.E. THEY CANNOT BE DIRECTLY INTERPRETED IN TERMS OF THE ORIGINAL SLIDER RESPONSES' RANGE. INTERCEPT REFERS TO THE BASIC FACTOR COMBINATION (SOUND = SOUND 1, VIBRATION = SINE, ATTRIBUTE = *PREFERENCE*). THE EFFECTS OF THE OTHER LEVELS OF THE CATEGORICAL FACTORS AND THEIR INTERACTIONS ARE ADDITIVE TO THE INTERCEPT. AN EFFECT IS CREDIBLE (MARKED WITH BOLD FONT) IF ITS CI DOES NOT CONTAIN ZERO

Parameter	Estimate	Q2.5%	Q97.5%
mu Intercept	0.25	0.04	0.47
mu sound2	-0.11	-0.31	0.09
mu vibrationAudio	-0.06	-0.25	0.13
mu vibrationNoise	-0.34	-0.55	-0.13
mu attributeEnjoy	0.11	-0.03	0.26
mu attributeExpr	0.13	-0.01	0.27
mu attributeControl	-0.01	-0.15	0.14
mu sound2:vibrationAudio	0.13	-0.11	0.36
mu sound2:vibrationNoise	0.25	0.01	0.48
phi Intercept	1.22	1.01	1.42
phi vibrationAudio	-0.30	-0.49	0.12
phi vibrationNoise	-0.32	-0.51	-0.13
phi sound2	0.31	0.16	0.47
phi attributeEnjoy	0.23	0.02	0.44
phi attributeExpr	0.45	0.24	0.67
phi attributeControl	0.41	0.20	0.63
zoi Intercept	-1.10	-1.48	-0.75
zoi vibrationAudio	-0.43	-0.81	-0.04
zoi vibrationNoise	-0.57	-0.98	-0.17
zoi sound2	-0.35	-0.66	-0.03
zoi attributeEnjoy	-0.22	-0.63	0.18
zoi attributeExpr	-1.21	-1.76	-0.68
zoi attributeControl	-0.54	-0.97	-0.11
coi Intercept	1.09	0.34	1.88
coi vibrationAudio	-0.21	-1.05	0.59
coi vibrationNoise	-0.66	-1.51	0.17
coi sound2	-0.55	-1.22	0.13
coi attributeEnjoy	0.47	-0.38	1.35
coi attributeExpr	-0.19	-1.28	0.91
coi attributeControl	0.00	-0.89	0.91

where the mean of the beta distribution is modeled by sound, vibration type, their interaction , and attribute. In addition, variation by subject (sID) is allowed for sound and vibration, as described by the last term in the model. The models for the precision (ϕ) and zero-one-inflation parameters (zoi, coi) are set to depend on vibration type, sound, and attribute without interactions.

Estimates for the beta distribution means and their corresponding 95% Credible Intervals are presented in Fig. 7. On average, the vibrating setups were preferred to their non-vibrating versions: all mean estimates but one are above 0.50 (the point of perceived equality) as well as most of the respective credible intervals. As seen in Fig. 7 (b), the presence of vibrations had the strongest positive effect on *Expression* and *Enjoyment*. Figure 7 (a) shows that the effects of vibrations were somewhat stronger on Sound 1 than on Sound 2. In both cases, Audio and Sine vibration had a positive effect, while Noise did not significantly increase the perceived quality of the interface.

The marginal effect of sound type was not credible in the ZOIB model. In combination with Noise, however, Sound 2 had a slight but credible positive effect. Parameter estimates and their 95% Credible Intervals are given in Table I.

TABLE II
ESTIMATED μ PARAMETERS FROM THE ZOIB FIT (ON ORIGINAL RESPONSE SCALE) FOR THE MARGINAL EFFECTS OF SOUND AND VIBRATION (ATTRIBUTE = *PREFERENCE*). N=29: ALL SUBJECTS; N=19: CONSISTENT SUBJECTS

Sound	Vibration	Estimate (N=29)	Estimate (N=19)
1	Sine	0.563	0.604
1	Audio	0.548	0.576
1	Noise	0.480	0.466
2	Sine	0.536	0.558
2	Audio	0.552	0.550
2	Noise	0.512	0.493

The ZOIB model parameters are modeled on transformed scales using so-called *link functions* in the same way as in general linear models. The parameter values in Table I are described on the scale of their respective link functions: The mean and the inflation parameters were logit-transformed[10], while precision was modeled on the log scale. Therefore these values cannot be directly interpreted in terms of probability or the original slider response range. Conversely, Fig. 7 and Tables II and III present the estimated effects transformed back to the slider response scale [0,1].

In terms of distribution means, Audio vibration was not significantly different from Sine vibration, while Noise vibration was rated credibly lower. However, both Audio and Noise vibrations had a significant effect on the precision parameter (ϕ) of the beta distribution, as well as on the zero-inflation parameter (zoi), suggesting that even if the beta distribution means of the Audio and Sine vibration are not credibly different, the shapes of the respective distributions and proportions of $\{0, 1\}$ ratings may differ.

Sound type had a credible effect on the mean parameter (μ) only in combination with Noise vibration. However, the precision (ϕ) and inflation (zoi) estimates were credibly non-zero even for sound type, suggesting differences in the respective distribution shape.

Although we cannot conclude that any of the attributes have a clearly significant effect on the μ parameter, the credible intervals of both *Expression* and *Enjoyment* are almost entirely above zero in Table I, suggesting that their positive effect on ratings is in fact rather credible. In addition, all attributes have a credible effect on either the precision or inflation parameters, again suggesting differences in the respective distribution shape. Furthermore, all four attributes were somewhat positively correlated: *Preference*, which was rated in isolation from the other attributes, had the strongest correlation with *Enjoyment* ($r = 0.37$).

2) Consistency: Response consistency across repetitions was evaluated by modeling participants' first and second round responses by linear regression. The regression coefficient for the model containing all factor combinations was 0.32 ($p < 0.001$***), indicating that in general participants behave consistently (i.e., they preferred the same vibrating or non-vibrating setup twice across repetitions). Consistency was generally higher for Sound 1 (Sine: 0.32***, Audio: 0.28***, Noise 0.56***) than for Sound 2 (Sine: 0.21**, Audio: 0.35***, Noise 0.16*).

TABLE III
ESTIMATED μ PARAMETERS FROM THE ZOIB FIT (ON ORIGINAL RESPONSE SCALE) FOR THE MARGINAL EFFECTS OF ATTRIBUTE (SOUND = SOUND 1, VIBRATION = SINE). N=29: ALL SUBJECTS; N=19: CONSISTENT SUBJECTS

Attribute	Estimate N=29	Estimate N=19
Preference	0.563	0.604
Control	0.562	0.594
Expression	0.594	0.645
Enjoy	0.591	0.628

Consistency of individual participants was assessed in a similar manner. It was observed that ten of them often preferred once the vibrating and once the non-vibrating setup in the same factor combination, resulting in regression coefficients ≤ 0 (mean coefficient over the N=10 subjects $=-0.19$). The remaining subjects (N=19) instead gave consistent ratings (mean coefficient 0.53). Interestingly, the inconsistent group (N=10) spent noticeably less time with the tasks than the reliable group (N=19): the median length of their gestural data logs (see III-B3) was only 62% of that of the consistent group.

Considering the free playing task, inconsistent ratings could be due to different playing behavior between the two repetitions rather than task difficulty. For this reason, we included all N=29 participants in the following gesture trajectory analysis. Nevertheless, in order to estimate the effect of the inconsistent participants, we re-run the ZOIB model including only the N=19 consistent subjects and finding that the main result was similar to the the full dataset: only vibration type had a significant effect. However, this way the effect is somewhat larger, as the mean estimates for vibration types Sine and Audio (with Sound 1) slightly increase, while that for Noise decreases (see Table II). Also in this case *Expression* is the highest rated attribute; its marginal mean estimate increases from to 0.59 to 0.64 (see Table III).

3) Gestural Data: Trajectories (x, y, z) performed on the HSoundplane by the participants were recorded, logging also the respective configuration (i.e., sound and vibration type). Although the present free playing task did not allow for an in-depth analysis of gestures in presence and absence of vibrations, trajectories were compared in terms of their general characteristics.

Figure 8 presents estimated density distributions for the x, y, and z data in presence and absence of vibration. The peaks in the x and y dimensions clearly follow the gridded configuration of the interface (chromatic subdivision). In terms of density distributions, vibrotactile feedback did not cause systematic differences in playing behavior. This was also generally true for the various factor combinations.

Even if participants did not play all that differently with and without vibrations, could their playing style correlate to their subjective ratings? Such connections were analysed as follows. At each factor combination, participants were classified into two groups according to their preference for either the vibrating or the non-vibrating setup (i.e., median *Preference* and *Expression* ratings at that factor combination > 0.5 or

---

[10] The logit function maps values from $p \in [0, 1]$ to $x \in [-\infty, \infty]$ according to $x = \log\left(\frac{p}{1-p}\right)$; the inverse mapping is given by the logistic function $p = \frac{1}{1+e^{-x}}$.

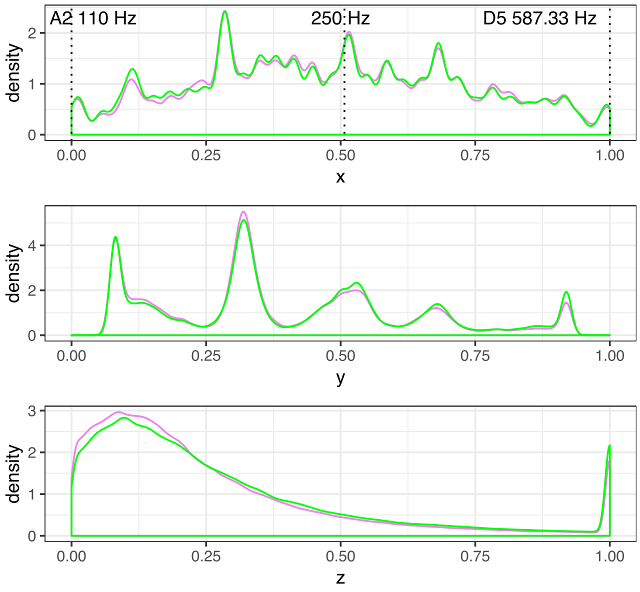


Fig. 8. Estimated density distributions of x, y, and z coordinates normalized in the [0 1] range, in presence (green) and absence (violet) of vibrations.

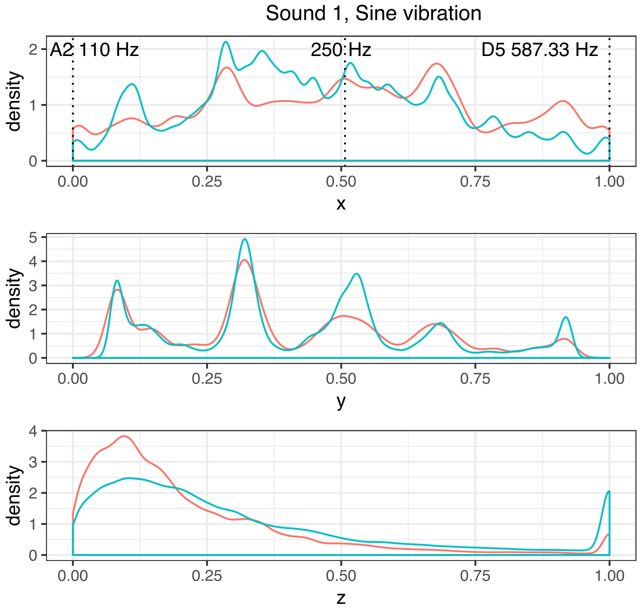


Fig. 9. Estimated density distributions of x, y, and z coordinates normalized in the [0 1] range for the configuration (Sound 1, Sine vibration). Blue: subjects who preferred the vibrating setup (N=22); red: subjects who preferred the non-vibrating setup (N=7).

≤ 0.5, respectively). Density distributions of x, y, and z coordinates in presence of vibration were then estimated and compared between “positive” and “negative” groups. The results are seen for the most preferred setup (Sound 1, Sine vibration) in Fig. 9, showing some notable differences: Concerning the x-axis (pitch control), the positive group spent more time in the lower range than the negative group — their 50-th percentiles (Q2) are respectively located at x=0.44 and x=0.51; Along the y-axis (controlling vibrato intensity), the positive group occupied a wider range than the negative group (Q2: y=0.42 and y=0.35, respectively); In the z dimension (controlling sound level and filter cutoff frequency), the negative group pressed more lightly than the positive group (Q2: z=0.15 and z=0.23, respectively). For Sound 1 with Noise vibration (Fig. 10) – that is the least preferred setup – such differences were less apparent. Along the x-axis, the differences in pitch are smaller (Q2: x=0.45 and x=0.49 for positive and negative groups, respectively). Concerning the y-axis data, contrary to Sine vibrations here the negative group used slightly more vibrato (Q2: y=0.37 and y=0.44). There was no noticeable difference in pressing force (Q2: z=0.22 and z=0.22).

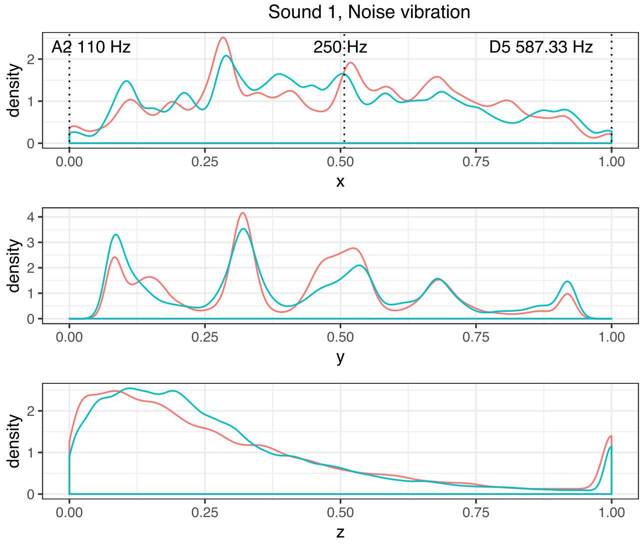


Fig. 10. Estimated density distributions of x, y, and z coordinates normalized in the [0 1] range for the configuration (Sound 1, Noise vibration). Blue: subjects who preferred the vibrating setup (N=14); red: subjects who preferred the non-vibrating setup (N=15).

IV. DISCUSSION

The main findings of this study are shown in Fig. 7: i) although not large, the measured effect of Sine or Audio vibration was appreciably positive; ii) Noise vibration did not enhance the subjective quality of the interface as compared to the non-vibrating condition; iii) vibrotactile feedback especially increased the perceived expressiveness of the interface and the enjoyment of playing. In line with the results of previous related studies, which assessed qualitative effects of haptic feedback in digital musical interfaces [19], [20], [35], [36], it can be concluded that vibrotactile feedback has the potential to improve the perceived quality of an interactive surface for musical expression, and the related playing experience.

As appears from Fig. 7 (a), a more marked effect was found when vibration was more similar to the sonic feedback and consistent with the user’s gesture: Indeed Sine and Audio vibration follow the pitch of the produced sound and their intensity can be controlled by pressure. Conversely, Noise vibration – offering fixed amplitude, independent of the input gesture, and flat spectrum – was rated lowest among the vibrating setups. Noise vibration resulted in slightly better ratings when Sound 2 was used as compared to Sound 1: Again, that was likely because vibrotactile feedback is consistent, at least partially, with the noise-like sonic feedback produced for high y-axis values. Interestingly, no credible difference in the globally positive effect was found between Sine and Audio vibration. This may be at least partially explained by a masking effect taking place in the tactile domain towards higher frequencies, thus impairing waveform discrimination [28]. However such phenomenon seems not to apply to markedly different signals [37]. In this regard, our informal testing revealed that Sine and Audio vibration were virtually indistinguishable, especially when Sound 1 (modulated sawtooth waveform) was selected. From a practical perspective, this fact may result in a technical advantage when implementing haptic musical interfaces:.[11] Indeed, it is rather trivial to generate sinusoidal signals at audio-rate on low-cost embedded systems, and in this way the effectiveness of vibrotactile feedback would be always optimized for rendition (e.g., as reported in II-C, the employed piezo actuators can efficiently and accurately reproduce sinusoidal signals in the $110-587.33\,\mathrm{Hz}$ range) and independent of the chosen auditory feedback (e.g., signals whose spectral energy is mainly above the range of tactile sensitivity or below the cutoff of the system would not be suitable); also, having to feed external multi-channel audio signals back into the interface would be impractical.

Despite the credible effects observed in the test population, in III-B2 we reported a number of inconsistent responses. As the participants were highly skilled musicians, we do not believe that the task was too difficult. However, as they were not screened for individual vibrotactile sensitivity, it is possible that they did not feel vibrations equally strong. Additionally, we observed some trade-off between execution speed and response reliability, which is typical of decision-making tasks [39]. On top of that, we argue that rating inconsistency may be linked to the varying perceived vibration strength and audio-tactile congruence, depending on where and how the participants were playing over the interface's surface. This aspect is examined in more detail further below. Indeed, vibrotactile intensity perception is affected by vibration amplitude (obviously), spectral content (with a peak in the $200-300\,\mathrm{Hz}$ range [29]), and the exerted pressing force [2]; also, varying degrees of spectral and temporal similarity between auditory and vibratory feedback may result either in crossmodal perceptual integration or interference [40]. However, we specifically chose a free playing task in order to measure the effect of vibrotactile feedback on various aspects of the playing experience.

With regard to the coherence of specific audio-tactile combinations, although Noise vibration resulted in very uniform ratings when associated with Sound 1, it produced the lowest rating consistency with Sound 2. While this was obviously affected by the general tendency of ten participants towards inconsistent ratings, one may also consider the varying degree of similarity between Sound 2 and Noise vibration: at the upper range of the y coordinate Sound 2 was noise-like, while for lower y values it was increasingly sinusoidal; inconsistency might follow from having played once mostly at high y and once mostly at low y. Conversely, Sound 1 retained the same degree of (dis)similarity with Noise vibration, independent of the playing position/style. Overall, the noticed inconsistency of responses sets a future challenge for screening the participants and controlling the playing task.

Concerning the gestural data reported in III-B3, several observations can be made, especially with regard to the positive group who preferred the combination of Sound 1 and Sine vibration (see Fig. 9): i) Subjects in this group spent more time at a pitch range (i.e., x-axis values) that results in strong vibration sensation [29]; ii) They used a wider y-axis range, resulting in more variation in the produced sound. iii) On average they applied a higher pressing force (i.e., z-axis values), again producing stronger tactile sensation [2]. With regard to the combination of Noise vibration and Sound 1, the average coordinate values reported at the end of III-B3 seem to suggest that, while playing vibrato (i.e., at high y-axis values), such vibration was perceived as unpleasant. Did subjects give positive ratings because they happened to be playing in a way that emphasized the beneficial effect of vibrations? Or conversely, did vibrations motivate them to play that way and give a positive rating? In order to answer these questions and better highlight correlations between ratings, trajectories, vibration type, and vibration sensation, a more controlled setting is required, where subjects are not allowed to play and switch vibration on and off freely.

During debriefing almost all participants stated to having enjoyed very much playing on the HSoundplane, especially when vibrotactile feedback was present. However, a few of them reported that vibration was sometimes too strong, suggesting that adjustable intensity may be a crucial design parameter for improving comfort. Indeed, a weak negative correlation ($r = -0.12$) was observed in the negative group between participants' mean pressing force (z-axis data) and ratings, while in the positive group this correlation was weakly positive ($r = 0.09$). We plan to further investigate this aspect in a follow-up experiment in which different vibration levels will be tested.

V. CONCLUSION

A subjective assessment was conducted using an original musical interface which offers a large surface sensing multi-touch position and force, and yields multi-point advanced vibrotactile feedback. The goal was to evaluate how the presence and type of vibration influence the playing experience and the perceived quality of the device. Vibration signals providing spectral and amplitude dynamics cues consistent with the auditory feedback and/or the user's gestures were clearly preferred to the respective non-vibrating setup. Conversely, vibration with fixed flat spectrum and amplitude – i.e., independent from the pitch and dynamics of the played tones – had null or even slightly detrimental effects.

---

[11] Musical interfaces, such as the HSoundplane, are controllers for external sound generation or processing (e.g., a mixer). Conversely, digital musical instruments (DMIs) are self-contained devices formed by an interface and a sound synthesis/processing engine [38].

More complex strategies were recently proposed in the literature to design tactile signals based on the auditory feedback so as to enrich the music listening experience [41], [42] — these may be tested in a future experiment aiming to further highlight the positive effect of vibrotactile feedback.

The rendition of multi-point, wide-band vibrotactile feedback on an interactive surface can pose serious technical challenges, especially in the perspective of industrial realization. Nevertheless, based on the reported results we suggest that the design of future interactive surfaces for musical expression – and by extension next generation digital musical interfaces in general – should take into consideration the addition of advanced vibrotactile feedback. This would enable the re-establishment of a consistent physical exchange between musicians and their digital musical devices – similar to what naturally found on acoustic musical instruments, where the source of sound and vibration coincides – with the demonstrated potential to enhance the playing experience and the perceived quality of the interface. Indeed several participants in our study reported to be impressed with the novelty and "aliveness" of the HSoundplane, as opposed to their experience with existing digital musical interfaces.

Ultimately, it is yet to be seen if and how such subjective enhancements may be reflected in the quality of playing, and musical performance altogether. Making objective measurements of these aesthetic aspects however poses a major research challenge, and the present work only scratched the surface in this direction. Instead, this will be the main object of a follow-up experiment currently in the works.

ACKNOWLEDGMENT

The authors would like to thank Martin Fröhlich, and Thomas Peter at the Zurich University of the Arts for helping with the implementation of the HSoundplane's hardware and software systems.

REFERENCES

[1] Y. Shao, V. Hayward, and Y. Visell, "Spatial patterns of cutaneous vibration during whole-hand haptic interactions," in *Proc. National Academy Sci. United States Amer.*, vol. 113, no. 15, Mar. 2016, pp. 4188–4193.

[2] S. Papetti, H. Jarvelainen, B. L. Giordano, S. Schiesser, and M. Frohlich, "Vibrotactile sensitivity in active touch: Effect of pressing force," *IEEE Trans. Haptics*, vol. 10, no. 1, pp. 113–122, Jan. 2017.

[3] C. Nam and D. Shin, "Force-touch measurement methodology based on user experience," *Int. J. Distrib. Sensor Netw.*, vol. 14, no. 4, 2018.

[4] C. Hudin, J. Lozada, and V. Hayward, "Localized tactile feedback on a transparent surface through time-reversal wave focusing." IEEE Trans. Haptics, vol. 8, no. 2, pp. 188–198, Apr.–Jun. 2015.

[5] J. Mullenbach, M. Peshkin, and J. E. Colgate, "eShiver: Force feedback on fingertips through oscillatory motion of an electroadhesive surface," *IEEE Trans. Haptics*,vol. 10, no. 3, pp. 358-370, Jul.–Sept. 2017.

[6] S. Jordá, M. Kaltenbrunner, G. Geiger, and R. Bencina, "The reactable," in *Proc. Int. Comput. Music Conf.*, 2005, pp. 579–582.

[7] S. Papetti and C. Saitis, Eds., *Musical Haptics*, Berlin, Germany: Springer, 2018.

[8] S. Papetti and C. Saitis, "Musical haptics: Introduction," in *Musical Haptics*, S. Papetti and C. Saitis, Eds. Berlin, Germany: Springer, 2018, pp. 1–7.

[9] G.-M. Schmid, "Measuring musician's playing experience: Development of a questionnaire for the evaluation of musical interaction," in *Proc. Workshop Creativity Cogn. Conf. New Interfaces Musical Expression*, 2014.

[10] F. Fontana, S. Papetti, H. Järveläinen, F. Avanzini, and B. L. Giordano, "Perception of vibrotactile cues in musical performance," in *Musical Haptics*, S. Papetti and C. Saitis, Eds. Berlin, Germany: Springer International Publishing, 2018, pp. 49–72.

[11] C. Saitis, H. Järveläinen, and C. Fritz, "The role of haptic cues in musical instrument quality perception," in *Musical Haptics*, S. Papetti and C. Saitis, Eds. Berlin, Germany: Springer International Publishing, 2018, pp. 73–93.

[12] G. W. Young, D. Murphy, and J. Weeter, "A functional analysis of Haptic feedback in digital musical instrument interactions," in *Musical Haptics*, S. Papetti and C. Saitis, Eds. Berlin, Germany: Springer, 2018, pp. 95–122.

[13] F. Fontana, S. Papetti, H. Järveläinen, and F. Avanzini, "Detection of keyboard vibrations and effects on perceived piano quality," *J. Acoustical Soc. Amer.*, vol. 142, no. 5, pp. 2953–2967, Nov. 2017.

[14] M. Keane and G. Dodd, "Subjective assessment of upright piano key vibrations," *Acta Acustica United Acustica*, vol. 97, no. 4, pp. 708–713, jul 2011.

[15] A. Galembo and A. Askenfelt, "Quality assessment of musical instruments - Effects of multimodality," in *Proc. ESCOM Conf.*, 2003, pp. 441–444.

[16] I. Wollman, C. Fritz, and J. Poitevineau, "Influence of vibrotactile feedback on some perceptual features of violins." *J. Acoustical Soc. Amer.*, vol. 136, no. 2, pp. 910–921, Aug. 2014.

[17] I. Hwang, H. Son, and J. R. Kim, "AirPiano: Enhancing music playing experience in virtual reality with mid-air haptic feedback," in *Proc. IEEE World Haptics Conf.*, 2017, pp. 213–218.

[18] G. W. Young, D. Murphy, and J. Weeter, "A qualitative analysis of Haptic feedback in music focused exercises," in *New Interfaces for Musical Expression*. 2017, pp. 204–209.

[19] 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 *Sound and Music Comput.*, 2015.

[20] F. Kalantari, F. Berthaut, and L. Grisoni, "Enriching musical interaction on tactile feedback surfaces with programmable friction," in *Proc. Int. Symp. Comput. Music Multidisciplinary Res.*, 2017, pp 387–401.

[21] M. E. Altinsoy and S. Merchel, "Touchscreens and musical interaction," in *Musical Haptics*, S. Papetti and C. Saitis, Eds. Berlin, Germany: Springer, 2018, pp. 239–255.

[22] R. Jones, P. Driessen, A. Schloss, and G. Tzanetakis, "A force-sensitive surface for intimate control," in *Proc. Conf. New Interfaces Musical Expression*, 2009, pp. 236–241.

[23] J. Deber *et al.*, "Hammer time!: A low-cost, high precision, high accuracy tool to measure the latency of touchscreen devices," in *Proc. Conf. Human Factors in Comput. Syst.*, 2016, pp. 2857–2868.

[24] S. Papetti, M. Fröhlich, F. Fontana, S. Schiesser, and F. Avanzini, *Implementation and Characterization of Vibrotactile Interfaces*. Berlin, Germany: Springer, 2018, pp. 257–282.

[25] S. J. Bolanowski, G. a. Gescheider, R. T. Verrillo, and C. M. Checkosky, "Four channels mediate the mechanical aspects of touch." *J. Acoust. Soc. Amer.*, vol. 84, no. 5, pp. 1680–94, Nov. 1988.

[26] A. Farina, "Advancements in Impulse Response Measurements by Sine Sweeps," in *Audio Eng. Soc. Conv. 122. Audio Engineering Society*, 2007.

[27] Y. D. Pra, S. Papetti, F. Fontana, and M. Simonato, "An Open-Source Remote-Controllable Hardware Device for Realistic Finger-Pressing Simulation," 2020. [Online]. Available: https://eurohaptics2020.org/program/posters-demos-2/

[28] S. J. Bensmaïa and M. Hollins, "Complex tactile waveform discrimination," *J. Acoust. Soc. Amer.*, vol. 108, no. 3, pp. 1236–1245, 2000.

[29] R. T. Verrillo, "Vibration sensation in humans," *Music Perception*, vol. 9, no. 3, pp. 281–302, 1992.

[30] J. Ramsagaard, T. Worch, and N. Zacharov, "Sensory evaluation methods for sound," in *Sensory Evaluation of Sound*, N. Zacharov, Ed, Boca Raton, FL, USA: CRC Press, New York, NY, USA: Taylor & Francis Group, 2018.

[31] J. K. Kurschke, Doing Bayesian data analysis - A *tutorial with R. JAGS and Stan*, 2nd ed. New York, NY, USA: Academic Press, 2014.

[32] F. Liu and Y. Kong, "Zoib: An R package for Bayesian inference for beta regression and Zero/one inflated beta regression," *R J.*, vol. 7, no. 2, pp. 34–51, 2015.

[33] R. Ospina and S. L. P. Ferrari, "Inflated beta distributions," *Stat. Pap.*, vol. 51, pp. 111–126, 2010.

[34] P.-C. Bürkner, "Brms: An R package for Bayesian multilevel models using stan," *J. Stat. Software, Articles*, vol. 80, no. 1, pp. 1–28, 2017.

[35] O. Tache, J.-l. Florens, S. Sinclair, and M. M. Wanderley, "Exploring audio and tactile qualities of instrumentality with bowed string simulations," in *Proc. NIME 2012-12th Int. Conf. New Interfaces Musical Expression*. 2012.

[36] M. T. Marshall and M. M. Wanderley, "Examining the effects of embedded vibrotactile feedback on the feel of a digital musical instrument," *New Interfaces Musical Expression*, Jun., 2011, pp. 399–404.

[37] F. A. Russo, P. Ammirante, and D. I. Fels, "Vibrotactile discrimination of musical timbre." *J. Exp. Psychol. Hum. Percept. Perform.*, vol. 38, no. 4, Aug. 2012.

[38] E. R. Miranda and M. M. Wanderley, *New Digital Musical Instruments: Control and Interaction Beyond the Keyboard*, 2006.

[39] R. P. Heitz, "The speed-accuracy tradeoff: History, physiology, methodology, and behavior," *Front. Neurosci.*, vol. 8, 2014.

[40] J. M. Yau, J. B. Olenczak, A. I. Weber, J. F. Dammann, and S. J. Bensmaia, "Pitch and loudness interactions between audition and touch," *J. Acoust. Soc. Amer.*, vol. 129, no. 4, pp. 2525–2525, Apr. 2011.

[41] S. Merchel and M. E. Altinsoy, "Auditory-tactile experience of music," in *Musical Haptics*, S. Papetti and C. Saitis, Eds. Berlin, Germany: Springer, 2018, pp. 123–148.

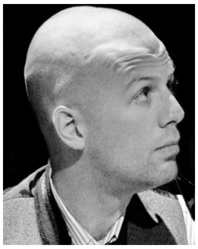
[42] R. Okazaki, H. Kuribayashi, and H. Kajimoto, "The effect of frequency shifting on audio-tactile conversion for enriching musical experience," *Lect. Notes Electr. Eng.*, 2015, vol. 277, pp. 45–51.

Stefano Papetti (Member, IEEE) received the M.Sc. degree (Laurea) in computer engineering from the University of Padua, Padua, Italy, in 2006 and the Ph.D. degree in computer science from the University of Verona, Verona, Italy, in 2010.

He is currently a Research Associate with the Institute for Computer Music and Sound Technology, Zurich University of the Arts, Zürich, Switzerland, where he leads the research on haptic musical interaction. His current research interests include the design and evaluation of haptic musical interfaces, and models and applications for interactive sound synthesis.

Hanna Järveläinen received the Bachelor of Music degree in music theory and composition from the Sibelius Academy, Helsinki University of the Arts, Helsinki, Finland, in 2006, and the M.Sc. and Ph.D. degrees in electronic and communications engineering from the Aalto University (then Helsinki University of Technology), Espoo, Finland, in 1997 and 2003, respectively.

She is currently a Research Associate with the Institute for Computer Music and Sound Technology, Zurich University of the Arts. Her research interests include psychoacoustics and multisensory perception and action in music. She studied early music with the Schola Cantorum Basiliensis, Basel, Switzerland, and is active as a Singer.



Sébastien Schiesser received the M.Sc. degree in microtechnology from the Swiss Federal Institute of Technology, EPFL, Lausanne, Switzerland, and the Performance degree in saxophone from the Zurich University of the Arts, Zürich, Switzerland.

He is currently a Freelance Musician and a Research Associate with the Institute for Computer Music and Sound Technology, Zurich University of the Arts, where he designs electronics and develops new musical interfaces.