RENDERING AND SUBJECTIVE EVALUATION OF REAL VS. SYNTHETIC VIBROTACTILE CUES ON A DIGITAL PIANO KEYBOARD

Federico Fontana
Università di Udine
Dept. of Mathematics and Computer Science
via delle Scienze 206, Udine 33100, Italy
`federico.fontana@uniud.it`

Federico Avanzini
Università di Padova
Dept. of Information Engineering
Via G. Gradenigo 6/b, Padova 35131, Italy
`avanzini@dei.unipd.it`

Hanna Järveläinen, Stefano Papetti
Zürcher Hochschule der Künste
Inst. for Computer Music and Sound Technology
Pfingstweidstrasse 96, Zurich 8031, Switzerland
`name.surname@zhdk.ch`

Giorgio Klauer, Lorenzo Malavolta
Conservatorio di Musica "Cesare Pollini"
Via Eremitani 18, Padova 35121, Italy
`name.surname@gmail.com`

ABSTRACT

The perceived properties of a digital piano keyboard were studied in two experiments involving different types of vibrotactile cues in connection with sonic feedback. The first experiment implemented a free playing task in which subjects had to rate the perceived quality of the instrument according to five attributes: Dynamic control, Richness, Engagement, Naturalness, and General preference. The second experiment measured performance in timing and dynamic control in a scale playing task. While the vibrating condition was preferred over the standard non-vibrating setup in terms of perceived quality, no significant differences were observed in timing and dynamics accuracy. Overall, these results must be considered preliminary to an extension of the experiment involving repeated measurements with more subjects.

1. INTRODUCTION

Research on musical haptic perception is constantly growing, aiming at connecting measurable effects originated by tactile properties of an instrument to subjective preference judgments and, ultimately, to the musician's experience and specific aspects of his or her performance. Such research considers both traditional instruments such as pianos and violins [1,2], and extends to augmentations spanning the broader area of new instrument design with applications to musical interaction and education [3,4].

Specifically concerning the piano, the reproduction of the tactile properties of the keyboard has been first approached from a kinematic perspective with the aim of reproducing the mechanical response of the keys [5, 6], also in light of experiments emphasizing the sensitivity of pianists to the keyboard mechanics [7]. Only recently, and in parallel

to industrial outcomes [8], did researchers start to analyze the role of vibrotactile feedback as a potential conveyor of salient cues: an early attempt claimed possible qualitative relevance of these cues [9]. Later along the same line, ground for a substantial step forward was set when some of the present authors not only found significant sensitivity to such cues [10], but also hypothesized that pianists are sensitive to key vibrations also when their amplitude is below the standard subjective thresholds originally estimated by stimulating subjects' fingertips with purely sinusoidal stimuli [11].

This conclusion gives rise to an interesting discussion, since it contradicts previous experiments [1] only apparently. Indisputably, those experiments did not take into account the complex perceptual effects due to vibrotactile temporal, spatial and spectral summations resulting from playing single or multiple keys. More importantly they did not address the issue of interactivity, reflecting an inherent lack of (also non-musical) studies addressing vibrotactile perception under active touch. Especially for this reason, authors of this paper have recently studied vibrotactile sensitivity measured under this condition, obtaining thresholds that are significantly lower than what reported in the previous literature [12].

In light of such unexpected differences found in the pianists' sensitivity thresholds, this work focuses on the ability of subjects to make a distinction in perceived *quality* between different types of vibrotactile feedback. In other words, we hypothesize that pianists appreciate the reproduction of real as opposed to simplified synthetic key vibrations. The experiment required to disassemble a digital piano keyboard, and instrument it so as to convey vibratory signals to the user; then, to record key vibrations on an acoustic piano and to synthesize simplified counterparts, which were organized in two respective sample banks.

Two experiments were planned making use of this setup: One studying subjective quality perception, and one measuring timing and dynamic performance. Results show that the setups augmented with vibrations were generally preferred over the non-vibrating standards, with a slight pref-

Figure 1. Experimental setup.

Figure 2. One of the transducers used to convey vibration at the keyboard.

erence towards amplified vibrations as opposed to vibrations of realistic amplitude. On the other hand, no effect was observed on timing or dynamics accuracy in the performance experiment, suggesting either that the difference is more of a subjective nature, or that vibrotactile feedback is not relevant for the specific performative task considered here. However, in the present pilot experiment, low concordance was observed between subjects, which suggests that intra- and inter-individual consistency is likely an important issue. As a future task, the experimental design needs some revision until more significant conclusions can be claimed.

2. SETUP

The keyboard of a Viscount Galileo VP-91 digital piano was detached from its metal casing, containing also the electric and electronic hardware, and then screwed to a thick plywood board (see Fig. 1).

Two Clark Synthesis TST239 Silver Tactile Transducers were attached to the bottom of the wooden board as shown in Fig. 2, respectively in correspondence of the lower and middle octaves, in this way enabling to convey vibrations at the most relevant areas of the keyboard [10]. Once equipped in this way, the keyboard was laid on a X-shaped keyboard stand, interposing foam rubber at the contact points.

The transducers were driven by a Yamaha P2700 amplifier in dual mono configuration, fed with a monophonic signal. The input was provided by a RME Fireface 800 audio interface communicating with an Apple MacBook Pro via Firewire. Sound and vibrotactile feedback were generated via software using Reaper 4 digital audio workstation,[1] which hosted the following plug-ins: the Pianoteq 4.5 physical modelling piano was used to synthesize audio feedback, delivered to the performer via headphones; the Native Instruments Kontakt 5 sampler[2] in series with MeldaProduction MEqualizer parametric equalizer[3] were used for vibration playback. The piano synthesizer was

configured to match the sound of the grand piano used for recording vibration samples, as described below.

A schematic of the setup is shown in Fig. 3. The computer was also used to conduct the tests and collect experimental data. For this, programs were implemented as patches for the Pure Data real-time environment.[4] More details are given below in the description of each experiment.

2.1 Spectral equalization

Even if the setup was assembled in a way to avoid resonances due to nonlinearities, evidently the vibratory frequency response of the keyboard-plywood board system was not flat. Additionally, the transducers exhibit a prominent notch around 300 Hz. The overall frequency response of the transduction-transmission chain was measured in correspondence of all the A keys and led to an average magnitude spectrum that, once inverted, provided the spectral flattening equalization characteristics shown in Fig. 4. It can be noticed that the 300 Hz notch of the transducers is compensated along with resonances and anti-resonances of the mechanical system.

To avoid the generation of resonance peaks along the keyboard, we approximated this characteristics using the parametric equalizer plug-in, namely with a shelving filter providing a ramp climbing by 18 dB in the range [100–600] Hz, in series with a 2nd-order filter block approximating the peak around 180 Hz.

2.2 Vibration signals

Real vibration recordings were acquired at the keyboard of a Yamaha DC3 M4 Disklavier, using a Wilcoxon Research 736 piezoelectric accelerometer and iT100M Intelligent Transmitter connected to the audio interface. By triggering each of the 88 actuated keys of the Disklavier via MIDI control, vibration samples were recorded on every key at velocities 12, 23, 34, 45, 56, 67, 78, 89, 100,

---

[1] www.reaper.fm
[2] www.native-instruments.com
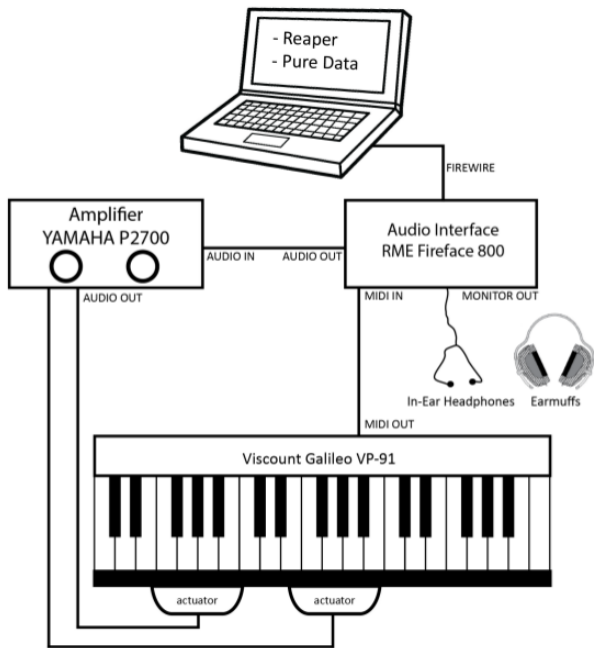[3] www.meldaproduction.com

[4] puredata.info
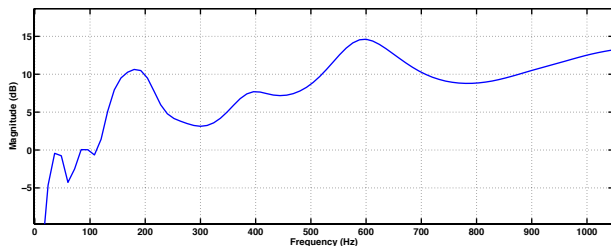
Figure 3. Schematic of the setup.

Figure 4. Spectral flattening: average equalization curve.

111. The accelerometer was secured to each measured key with Pongo. [5]

In addition to these recorded samples, a second bank of vibration signals was synthetically generated, with the purpose of reproducing the same amplitude envelope of the real signals while changing the spectral content ony. To this end, synthetic signals for each key and each velocity value were constructed as follows. First, white noise was was generated and then bandlimited in the range [20–500] Hz (corresponding to the vibrotactile bandwidth [11]). Then the noise was passed through a 2nd-order resonant filter centered at the fundamental frequency of the key. The resulting signal was modulated by the amplitude envelope of the corresponding recorded vibration sample, which in turn was estimated from the energy decay curve of the sample via the Schroeder integral [13]. Finally, the energy of the synthetic sample was equalized to that of the corresponding real sample.

The two sets of recorded and synthetic vibration samples were loaded into two distinct instances of the sampler plug-in, which managed their interpolation across velocities, based on the messages of MIDI note and key velocity coming from the digital keyboard.

---

[5] http://www.fila.it/en/pongo/history/

2.3 Key velocity calibration

The keys of the Disklavier and the Galileo digital piano have different response dynamics because of their mechanics and mass. Since pianists adapt their style in consequence of these differences, the digital keyboard had to be subjectively calibrated aiming at equalizing its dynamics with that of the Disklavier.

The keyboard response was set using the velocity calibration routine included in Pianoteq, which was performed by an experienced pianist first on the Disklavier and then on the digital keyboard. As expected, two fairly different velocity maps were obtained. Then, by making use of a MIDI filter plug-in in Reaper, each point of the digital keyboard velocity map was projected onto the corresponding point of the Disklavier velocity map. The resulting key velocity transfer characteristics was then independently checked by two more pianists, to validate its reliability and neutrality. In this way we ensured that when a pianist played the digital keyboard at a desired dynamics, the corresponding vibration samples recorded on the Disklavier would be triggered.

2.4 Loudness matching

As a final calibration step, the loudness of the piano synthesizer at the performer's ear was matched to that of the Disklavier grand piano. In order to do this, the sound produced by the A keys of the Disklavier at various velocities was recorded using a KEMAR mannequin positioned at the pianists location [10].

Then, additional measurements were taken with the KE-MAR mannequin now wearing the same equipment that would be later used by experimental subjects, i.e. a pair of Sennheiser CX 300-II earphones and, on top of them, a pair of 3M Peltor X5 ear-muffs. In this case, A notes generated at the corresponding velocities by the Pianoteq engine were played back. Finally, the loudness of the piano synthesizer was matched to that of the Disklavier, by using the volume mapping feature of Pianoteq, which allows one to set independently the volume of each key across the keyboard.

3. EXPERIMENTS

Eleven subjects participated in the experiment, five females and six males. Their average age was 26 years, and their average piano playing experience was 8 years after reaching conservatory level. Two of the subjects were jazz pianists, the rest played classical piano. All of them signed an informed consent form. A session including the two experiments lasted about one hour.

Audio-tactile stimuli were produced at runtime: the digital keyboard played by the participants sent MIDI messages to the computer, where the piano synthesizer plug-in generated the related sounds and, in parallel, the sampler plug-in played back the corresponding vibration samples then processed by the equalizer plug-in (see again Fig. 3).

Subjects wore earphones and ear-muffs on top of them, in the same fashion as the KEMAR mannequin did during the loudness matching procedure described above. In
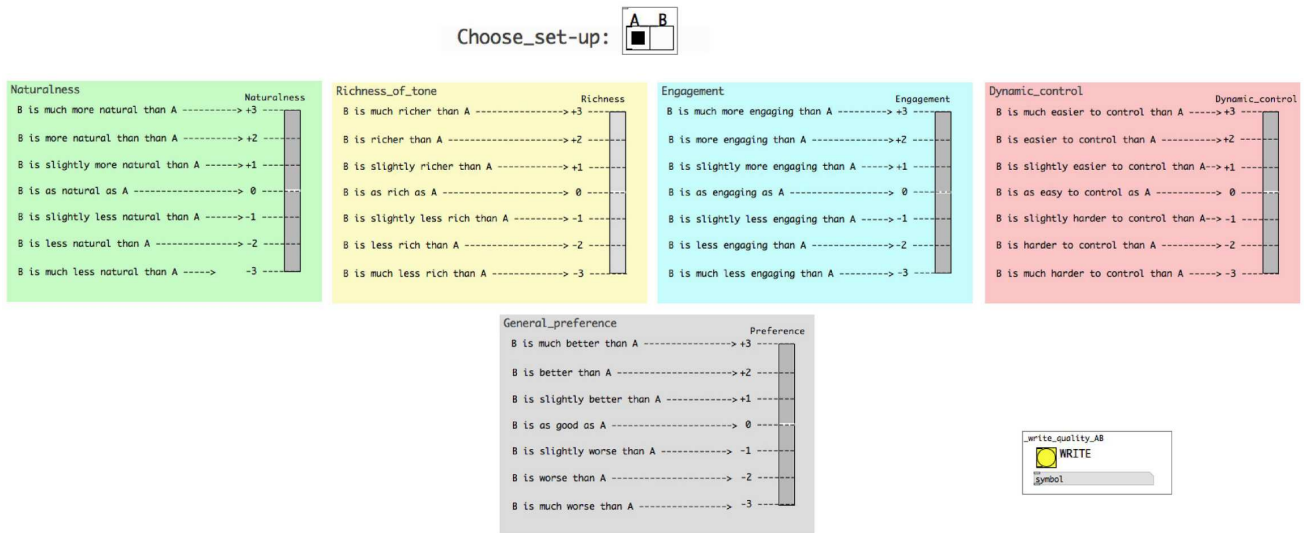
Figure 5. The graphical user interface used by participants to switch between conditions (A,B) and to rate the five attributes.

this way they were not exposed to the sound coming by air conduction from the transducers, as a by-product of their vibration.

3.1 Experiment 1: Quality

3.1.1 Stimuli and conditions

Three vibration conditions were assessed, always relative to a non-vibrating standard stimulus A:

B: recorded real vibrations;
C: recorded real vibrations with 9 dB boost;
D: synthetic vibrations.

Conversely, sound feedback was generated by the same piano synthesizer configuration throughout the experiment.

3.1.2 Design and procedure

The task was to play freely on the digital keyboard and assess the playing experience on five attribute rating scales: Dynamic control, Richness, Engagement, Naturalness, and General preference. The dynamics and range of playing were not restricted in any way.

Subjects could switch freely among setups α and β: Setup α was always the non-vibrating standard, while setup β was one of the three vibration conditions (B, C, D). The rating of β was given in comparison to α. The presentation order of the conditions was randomized. Also, participants were not aware of what could actually change in the different setups, and in particular they did not know that sound feedback would not be altered. The free playing time was 10 minutes per couple of conditions (A, B), and participants were allowed to rate the five attributes at any time during the session by means of a point & click graphical user interface (GUI), depicted in Fig. 5. In the end, each subject gave one rating in each attribute scale for each vibration condition.

Ratings were given on a continuous Comparison Category Rating scale (CCR), ranging from -3 to +3, which is widely used in subjective quality determination in communications technology (recommendation ITU-T P.800).

Subjects moved a slider on the continuous scale, to the position which best reflected their opinion. The scale had the following tick marks:

+3: "β much better than α"
+2: "β better than α"
+1: "β slightly better than α"
 0: "β equal to α"
-1: "β slightly worse than α"
-2: "β worse than α"
-3: "β much worse than α"

The five attributes were selected based on previous experiments with the Disklavier [10] and recent research on violin evaluation [2].

The GUI for the participants and the software for controlling the conditions and recording data were realized in Pure Data, running on a laptop placed at the subjects' reach.

## 3.2 Experiment 2: Timing and dynamic stability

The technical setup was the same as in Experiment 1. Additionally, a metronome sound at 120 BPM was delivered through the earphones.

Only conditions A and B were used in the test, alternating realistic vibrations with no vibrotactile feedback.

3.2.1 Design and procedure

Subjects were asked to play an ascending and then a descending D-major scale at pace with the metronome (every second beat), at a fixed given dynamics. Only the three leftmost octaves were considered, so as to maximize the tactile feedback, requiring the subjects to play with their left hand only. Each subject repeated the task for three dynamic levels (*pp*, *mf*, *ff*) three times each, in each condition (i.e., with and without vibrations), for a total of 18 randomized trials. MIDI data consisting of note ON, note length and key velocity messages were recorded across the test for subsequent analysis.

A program made with Pure Data was used to carry the test under the experimenter's supervision, and record MIDI data.
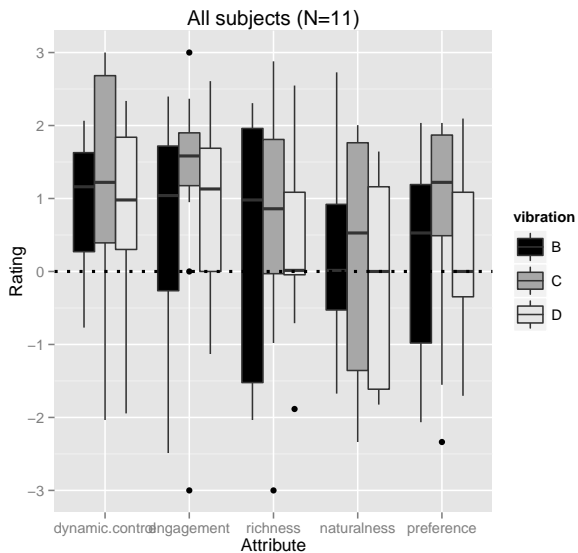
Figure 6. Results of the quality experiment. Boxplot presenting median and quartiles for each attribute scale and vibration condition.

Vibration	Dyn.	Rich.	Eng.	Nat.	Pref.
B	0.92	0.30	0.50	0.26	0.24
C	1.28	0.67	1.21	0.17	0.81
D	0.87	0.42	1.00	-0.23	0.29

Table 1. Mean ratings over all subjects for each attribute and vibration condition.

4. RESULTS

4.1 Perceived quality

Inter-individual consistency was assessed for each attribute scale by computing the Lin concordance correlations ρ_c for each pair of subjects [14]. The average ρ_c were 0.018 for general preference, 0.006 for dynamic control, -0.04 for richness, -0.02 for engagement, and -0.04 for naturalness. In all scales, a few subjects either agreed or disagreed almost completely and, due to this large variability, ρ_c was not significantly different from 0 for any of the scales ($t(54) < 0.77, p > 0.05$). The low concordance scores indicate a high degree of disagreement between subjects.

Responses were positively correlated between all attribute scales. The weakest correlation was observed between richness and dynamic control, (Spearman correlation $\rho_s = 0.18$), and the highest between general preference and engagement ($\rho_s = 0.75$). The partial correlations between general preference and the other attribute scales were as follows: $\rho_s = 0.39$ for dynamic control, $\rho_s = 0.72$ for richness, and $\rho_s = 0.57$ for naturalness.

Results are plotted in Fig. 6, and the mean ratings for each scale and vibration condition are given in Table 1. On average, each of the vibrating modes was preferred to the non-vibrating standard, the only exception being condition D for Naturalness. For conditions B and C Naturalness received faintly positive scores. The strongest preferences were for Dynamic range and Engagement. General preference and Richness had very similar mean scores though somewhat lower than Engagement and Dynamic control. Generally, C was the most preferred of the vibration conditions: it scored highest on four of the five scales, although B was considered the most natural. Interesting enough, B scored lowest in all other scales.

As the normality rule for Analysis of Variance was violated, a non-parametric Friedman test of differences among

repeated measures was conducted for the Preference ratings. It rendered a Chi-square value of 21.9 which was significant ($p < 0.05$), suggesting a significant difference between vibration conditions. However, Wilcoxon signed ranks tests, performed on the hypothesis that the median is positive, were insignificant for all conditions (B: $V = 37.5, p > 0.05$; C: $V = 41, p > 0.05$; D: $V = 28, p > 0.05$).
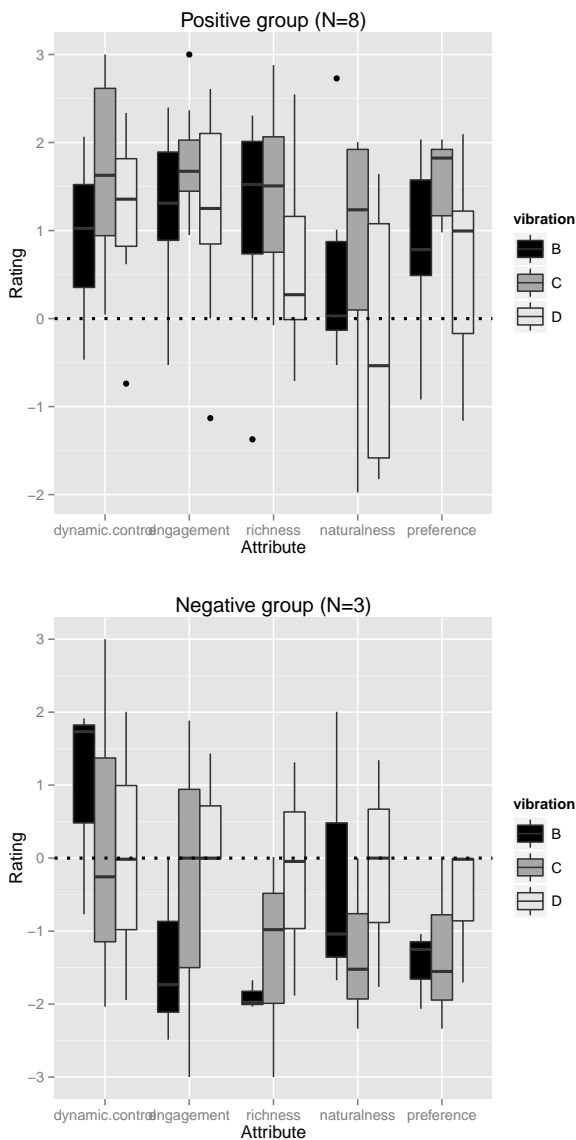
Heterogeneity was observed in the data, as might be expected due to the high degree of variability in the inter-individual agreement scores ρ_c. A k-means clustering algorithm was used to divide the subjects *a posteriori* into two classes according to their opinion on General preference. Eight subjects were classified into a "positive" group and the remaining three into a "negative" group. The results of the respective groups are presented in Fig. 7. A difference of opinion is evident: The median ratings for the most preferred setup C are nearly +2 in the positive group and -1.5 in the negative group for General preference. In the positive group, the median was > 0 in all cases except one (Naturalness, D), whereas in the negative group, the median was positive in only one case (Dynamic control, B).

4.2 Timing and dynamic stability

The hypothesis was that, if the subjects' timing and dynamic behaviour is affected by key vibrations, differences should be seen in means and standard deviations of key velocities and inter-onset intervals (IOI's).

Mean key velocities were computed for each subject as the average over the three repeated runs for each condition. Results are presented in Fig. 8. At the ff condition, subjects played just slightly louder with vibrations than without, while at the mf condition they played slightly softer in presence of vibrations. However, a repeated measures ANOVA was insignificant for both vibrations ($F(1, 2826) = 2.27, p > 0.05$) and the interaction between vibrations and dynamic condition ($F(2, 2826) = 0.83, p > 0.05$). No effect was observed either by studying the lowest octave alone, where the vibrations would be felt strongest. Nor was there a significant difference in the standard deviations between vibrations ON and OFF conditions (95% CI's obtained from paired t-tests (df=10) included $\mu(sd_B) - \mu(sd_A) = 0$ at all dynamic conditions).

IOIs were likewise stable across the two vibration conditions. Generally they were slightly more scattered at the pp dynamic, but no effect of vibrations was observed (see Fig. 9). Note durations were also stable across regardless of vibrations, suggesting that there was no significant difference in articulation or note overlap.
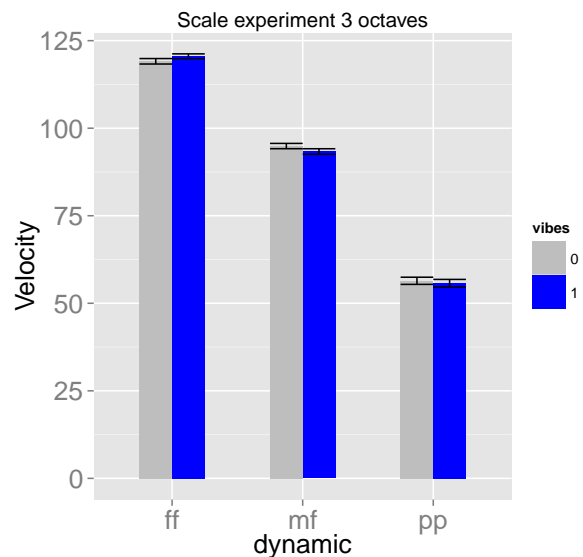
Positive group (N=8)

Negative group (N=3)

Figure 7. Quality results for the positive and negative groups.

Figure 8. Mean key velocities in Experiment 2, with 95% CI error bars as given in [15].

Figure 9. Boxplot of IOI's in the scale experiment.
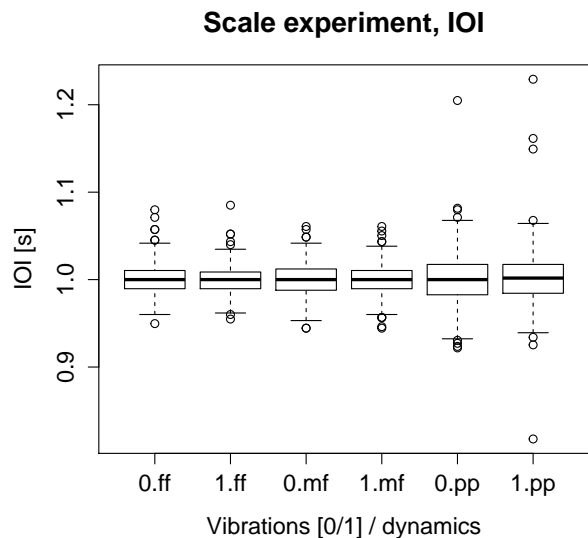
5. DISCUSSION AND CONCLUSIONS

It is concluded that key vibrations increase the perceived quality of a digital piano. Although the recorded vibrations were perceived as the most natural, amplified natural vibrations were overall preferred and received highest scores on all other scales as well. The other interesting outcome is that the vibrating setup was considered inferior to the non-vibrating standard only in Naturalness for synthetic vibrations. This suggests that pianists are indeed sensitive to the match between the auditory and vibrotactile feedback.

The attribute scales with the highest correlation to General preference were Engagement ($\rho_s = 0.75$) and Richness ($\rho_s = 0.72$). A similar result was obtained in a recent study on violin evaluation, where richness was significantly associated with preference [2].

The high degree of disagreement between subjects suggests that intra- and inter-individual consistency is an important issue in instrument evaluation experiments. Due to only one attribute rating per subject and condition, intra-

individual consistency could not be assessed in the present study and will be left for a future revision.

However, the heterogeneity in the data was similar across all attributes and conditions, making it hard to believe it was caused by inconsistency alone. Roughly two thirds of the subjects clearly preferred the vibrating setup, perhaps less rewarded by the synthetic vibrations, while the remaining one third had quite the opposite opinion. It is interesting that both the jazz pianists, having probably more experience of digital pianos than the classical pianists, were in the "negative" minority: would a vibrating digital keyboard be perceived as less pleasant than a neutral one, reflecting a preference of those pianists to the digital piano's traditional tactile response? In the next phase of the experiment, jazz and classical pianists will be studied in two *a priori* groups. Also, subjects will be asked to describe their opinion in a short qualitative interview.

No differences were observed in timing performance and dynamic stability. The task, three octaves of D-major scale in relatively slow pace, was probably easy to perform even without the possible aid of key vibrations. However, recent research shows that pianists use tactile information as a means of timing regulation [16, 17], even though the role of key vibrations remains unknown. There is also evidence that vibrotactile feedback helps force accuracy in finger pressing tasks [18, 19]. Whether vibrations caused by the currently depressed key(s) might help with velocity planning of the upcoming key press, is an interesting question that the present experiment cannot answer. Future experiments will investigate different performative tasks, in which the information provided by vibrotactile feedback is more salient than in the present one. As an example, a different task may involve repeated sustained chords where the player has to maintain the dynamics as constant as possible: in this case key vibrations and their time evolution are perceived more clearly and it may be hypotesized that they aid the control of dynamics.

It may well be that the effect of the key vibrations is purely subjective and simply makes playing the digital piano more engaging. However, subjects also gave high ratings for the perceived dynamic control for all vibrating setups. A future experiment will further investigate this effect on performance level.

Acknowledgments

6. REFERENCES

[1] A. Askenfelt and E. V. Jansson, "On vibration and finger touch in stringed instrument playing," *Music Perception*, vol. 9, no. 3, pp. 311–350, 1992.

[2] C. Saitis, B. L. Giordano, C. Fritz, and G. P. Scavone, "Perceptual evaluation of violins: A quantitative analysis of preference judgments by experienced players," *The Journal of the Acoustical Society of America*, vol. 132, no. 6, pp. 4002–4012, 2012. [Online]. Available: http://scitation.aip.org/content/asa/journal/jasa/132/6/10.1121/1.4765081

[3] M. Marshall and M. Wanderley, "Examining the effects of embedded vibrotactile feedback on the feel of a digital musical instrument," in *Proc. Int. Conf. on New Interfaces for Musical Expression (NIME)*, Oslo, Norway, May 30 - June 1 2011, pp. 399–404.

[4] M. Giordano and M. M. Wanderley, "Perceptual and technological issues in the design of vibrotactile-augmented interfaces for music technology and media," in *Haptic and Audio Interaction Design*, ser. Lecture Notes in Computer Science, I. Oakley and S. Brewster, Eds. Springer Berlin Heidelberg, 2013, vol. 7989, pp. 89–98. [Online]. Available: http://dx.doi.org/10.1007/978-3-642-41068-0_10

[5] C. Cadoz, L. Lisowski, and J.-L. Florens, "A Modular Feedback Keyboard Design," *Comput. Music J.*, vol. 14, no. 2, pp. 47–51, 1990.

[6] R. Oboe and G. De Poli, "A Multi-Instrument, Force-Feedback Keyboard," *Comput. Music J.*, vol. 30, no. 3, pp. 38–52, Sep. 2006.

[7] A. Galembo and A. Askenfelt, "Quality assessment of musical instruments - Effects of multimodality," in *Proc. 5th Triennial Conf. of the European Society for the Cognitive Sciences of Music (ESCOM5)*, Hannover, Germany, Sep. 8-13 2003.

[8] E. Guizzo, "Keyboard maestro," *IEEE Spectrum*, vol. 47, no. 2, pp. 32–33, Feb. 2010.

[9] F. Fontana, S. Papetti, M. Civolani, V. dal Bello, and B. Bank, "An exploration on the influence of vibrotactile cues during digital piano playing," in *Proc. Int. Conf. on Sound Music Computing (SMC2011)*, Padua, Italy, 2011, pp. 273–278.

[10] F. Fontana, F. Avanzini, H. Järveläinen, S. Papetti, F. Zanini, and V. Zanini, "Perception of interactive vibrotactile cues on the acoustic grand and upright piano," in *Proc. Joint ICMC/SMC Conf.*, 2014.

[11] T. Verrillo, "Vibrotactile thresholds measured at the finger," *Perception and Psychophysics*, vol. 9, no. 4, pp. 329–330, 1971.

[12] S. Papetti, H. Järveläinen, and G.-M. Schmid, "Vibrotactile sensitivity in active finger pressing," in *World Haptics Conf.*, 2015, to appear.

[13] M. R. Schroeder, "New method of measuring reverberation time," *J. Acoust. Soc. Am.*, vol. 37, no. 6, pp. 1187–1188, June 1965.

[14] L. Lin, "A concordance correlation coefficient to evaluate reproducibility," *Biometrics*, vol. 45, pp. 255–268, 1989.

[15] R. Morey, "Confidence intervals from normalized data: A correction to Cousineau (2005)," *Tutorial in Quantitative Methods for Psychology*, vol. 4, no. 2, pp. 61–64, 2008.

[16] W. Goebl and C. Palmer, "Tactile feedback and timing accuracy in piano performance," *Exp. Brain. Res.*, vol. 186, pp. 471–479, 2008.

[17] ——, "Finger motion in piano performance: Touch and tempo," in *International symposium for performance science*, 2009.

[18] H. Järveläinen, S. Papetti, S. Schiesser, and T. Grosshauser, "Audio-tactile feedback in musical gesture primitives: Finger pressing," in *Proceedings of the Sound and Music Computing Conference (SMC2013)*, Stockholm, 2013.

[19] T. Ahmaniemi, "Effect of dynamic vibrotactile feedback on the control of isometric finger force," *IEEE Trans. on Haptics*, 2012.