<table>
<thead>
<tr>
<th>Title</th>
<th>Haptics in music: the effects of vibrotactile stimulus in low frequency auditory difference detection tasks</th>
</tr>
</thead>
<tbody>
<tr>
<td>Authors</td>
<td>Young, Gareth W.; Murphy, David; Weeter, Jeffrey</td>
</tr>
<tr>
<td>Publication date</td>
<td>2016-12-29</td>
</tr>
<tr>
<td>Original Citation</td>
<td>Young, G. W., Murphy, D. and Weeter, J. (2016) ‘Haptics in music: the effects of vibrotactile stimulus in low frequency auditory difference detection tasks’. IEEE Transactions on Haptics, 10(1), pp. 135-139. doi:10.1109/TOH.2016.2646370</td>
</tr>
<tr>
<td>Type of publication</td>
<td>Article (peer-reviewed)</td>
</tr>
<tr>
<td>Link to publisher's version</td>
<td>10.1109/TOH.2016.2646370</td>
</tr>
<tr>
<td>Rights</td>
<td>© 2016, IEEE. Personal use of this material is permitted. Permission from IEEE must be obtained for all other uses, in any current or future media, including reprinting/republishing this material for advertising or promotional purposes, creating new collective works, for resale or redistribution to servers or lists, or reuse of any copyrighted component of this work in other works.</td>
</tr>
<tr>
<td>Download date</td>
<td>2023-11-14 17:10:42</td>
</tr>
<tr>
<td>Item downloaded from</td>
<td><a href="https://hdl.handle.net/10468/4419">https://hdl.handle.net/10468/4419</a></td>
</tr>
</tbody>
</table>
Haptics in Music: The Effects of Vibrotactile Stimulus in Low Frequency Auditory Difference Detection Tasks.

Gareth W. Young, Member, IEEE, Dave Murphy, and Jeffrey Weeter.

Abstract— We present an experiment that investigated the effect of vibrotactile stimulation in auditory pitch discrimination tasks. Extra-auditory information was expected to have some influence upon the frequency discrimination of auditory Just Noticeable Difference (JND) detection levels at 160 Hz. To measure this, the potential to correctly identified positive and negative frequency changes for two randomly divided groups was measured and then compared. The first group was given an audio only JND test and the second group was given the same test, but with additional vibrotactile stimulus delivered via a vibrating glove device. The results of the experiment suggest that in musical interactions involving the selection of specific pitches, or the detection of pitch variation, vibrotactile feedback may have some advantageous effect upon a musician’s ability to perceive changes when presented in synchrony with auditory stimulus.

Index Terms—Auditory (non-speech) feedback, Haptic I/O, Sound & Music Computing, Vibrotactile Feedback

1 INTRODUCTION

The manner in which auditory and haptic cues are integrated into musical performances with acoustic instruments are detailed in the findings of a number of studies, outlining the role therein of human senses beyond that of the auditory modality. Research has shown that the neural substrates of both the auditory and tactile systems are shared at a much lower level than previously understood. Furthermore, a cross-modal effect has been demonstrated in the tactile illusions that transpire from the modification of related audio stimuli; as seen in the “Parchment-skin illusion” [7]. Other auditory-tactile interactions have shown that tactile stimuli can influence auditory stimulus and vice-versa [8] [9] [10]. It can therefore be observed that auditory and haptic stimuli are capable of modifying or altering our perception of each when presented in unison. The experiment presented here primarily focuses on the detection of low frequency changes for both pure tone (sine wave) and more complex waveforms (saw and square waves); secondly, the musical ability of the participant was also considered.

2 BACKGROUND

Auditory and haptic feedback occurs in unison for most musical interactions that involve acoustic instruments, but tactile feedback itself rarely presents at a cognitive or decision making level. Research suggests that multimodal sensory cues are responsible for indirectly augmenting the auditory perception of music. This includes the influences of tactile and auditory feedback upon a performer, the performer’s understanding of the musical structure of a piece of music, and the portrayal of a score’s content. During a musical performance with an acoustic instrument, the control mechanisms of the performer rely on the multimodal feedback produced by the instrument. This feedback presents itself to the musician and they are then able to adjust and maneuver in response. Regardless of the manner of the interaction, via finger, hand, or lip placement, haptic feedback remains constant with auditory feedback. The transmission of vibrations to the performer in these interactions are an integral feature that directly relates to the design requirements of the acoustic instrument itself. However, Digital Musical Instruments (DMIs) are not restricted in this way and are capable of extending musical interactions beyond that of the acoustic experience.

Gillmeister and Eimer have previously highlighted the function of vibrotactile intensity enhancements when tactile stimulus is presented synchronously with auditory stimulus [8]. The interactions between these two stimuli produced mutual benefits and they followed principles of inverse effectiveness and the temporal rule of multisensory integration. From observing such effects in music, it is suggested that in the application of feedback in DMI design, vibrotactile information relating to the sound source being generated should be included. However, this application will ultimately depend on the musician’s ability to process this information in relation to the audio-visual feedback they are receiving and processing.

Physiological and Psychophysiological Studies

Auditory and tactile stimulation are received via physical mechanical pressure in the form of oscillations. These

1939-1412 (c) 2016 IEEE. Personal use is permitted, but republication/redistribution requires IEEE permission. See http://www.ieee.org/publications_standards/publications/rights/index.html for more information.
mechanical vibrations, within the cochlea and against the mechanoreceptors of the skin, activate neural impulses that are processed by the brain. The relationship between the neural processing of these two modalities of transduction have been previously discussed [16].

Both audio and tactile stimuli overlap within the same frequency range. However, limitations exist as the ear displays an increased sensitivity over the skin. Previous experiments with audio-frequency vibrotactile feedback have presented absolute thresholds of tactile detection for both simple and complex waveforms across a frequency range of 10 to 1000 Hz [17]. Within this range, the sub-thresholds of vibrotactile stimulus detection can be divided into distinct ranges pertaining to the frequencies that are cutaneously detectable and the waveforms being applied. This can also be seen in the absolute threshold of hearing, but over a much wider range. On average, the ear functions within an auditory range of approximately 20 to 20 kHz, while the tactile range of the skin encompasses a much narrower range of only 0.3 to 1 kHz. Within these overlapping ranges, vibrotactile information has been shown to stimulate the auditory cortex and tactile and auditory information may be perceived as interleaved signals [18] [19]. Furthermore, research has also shown that the auditory and vibrotactile systems combine whilst performing objective detection tasks regardless of the relative phase or the temporal synchrony of the stimulus [20]. This indicates that both neural pathways of the auditory and tactile systems combine through a common or related network.

Other studies have also shown evidence of interaction between the auditory and somatosensory systems at a multitude of stages within the human central nervous system [21]. Enhancements in auditory processing through the addition of tactile feedback have been observed and this elevates the response speeds to those of supra-threshold stimuli [22]. It has also been observed that improvements in the intensity perception of faint tones can be achieved with extra-auditory stimulus [23]. Other studies have indicated that the detection of a stimulus can be enhanced when simultaneously registering with two or more sensory modalities [24]. This demonstrates how the reinforcement of neural activity occurs when two modalities stimulate in near unison of time and place upon the body. More recent psychophysical studies have focused on the ability to discriminate between vibrotactile tonalities whilst being masked from an auditory source [16] [20] [25].

All of these findings suggest that the simultaneous combination of tactile and audio stimulation influences the perceptual discrimination of frequency within each sensory system. This is mainly attributed to the low-level integration of these two modalities in the cortical system. The relationships between the strengths of these two modes of stimulus should therefore directly relate to the individual psychophysical models constructed for the human senses. In this context, numerous examples of singular sensory modality interactions have been measured, but it is rarely the case in music that one singular sense is operating alone.

In music, many events seek to compete for combined sensory attention and a number of these are capable of stimulating the musician in several ways at once. Therefore, the study presented here has been designed to focus on audio frequency tactile stimulus as a supporting sensory input. Synchronous audio-tactile events are particularly ingrained in acoustic musical instrument performances where these combined perceptual aspects are inherently integrated. However, they are rarely included in commercial digital artifacts that are applied in the creation of music. It is therefore suggested that vibrotactile feedback may be applied in this context to improve the user’s perception of musical pitch.

3 Pitch Discrimination of Pure and Complex Waveforms

This experiment was designed to measure the pitch perception abilities of two groups for both pure and complex waveforms at a fundamental frequency of 160 Hz. Due to audio stimuli being the more appropriate sense applied in music, participants were instructed to focus only upon the auditory stimulus when making judgements. The context of this study was to investigate these relationships in a music domain; therefore, participants were asked to self-identify as musicians or as non-musician based upon a strict criterion.

Experiment Method

A two-alternative forced choice (2AFC) frequency discrimination task was used to measure the participants’ sensitivity to the applied stimuli. This technique is theoretically uncontaminated by fluctuations in criterion, but a response bias towards one or more observations may still exist [26]. Although extreme response strategies are rare in 2AFC tasks, the forced choice design does not guarantee the complete absence of bias. Therefore, to measure true sensitivity, bias was eliminated. This was achieved by calculating d’ from hit and false-alarm data and correcting the proportion of correct responses for bias, p(c).…

Participants

Participants were randomly divided into two groups by coin flip: Auditory-Only (heads) or Auditory-Tactile (tails). The participants then identified as being musician or non-musician based upon having been formally trained and actively performing in the last five years. The Auditory-Only group consisted of 10 males and 5 females aged 22 to 49 (MD = 28; SD = 8.79). In this group, 7 participants identified as musicians and 8 as non-musicians. The Audio-Tactile group consisted of 8 males and 7 females aged 21 to 40 (MD = 28; SD = 6.26). In this group, 10 participants self-identified as musicians and 5 as non-musicians. Participants self-reported as having no hearing difficulties or physical impairment.

Experiment Design

Participants were seated in a soundproofed room and asked to evaluate the relative pitch of two short audio samples. For the Auditory-Only group, dual mono audio
stimuli were delivered via Sennheiser HD215 headphones at 60 dB SPL (conversational speech at 1m). Participants were given the opportunity to adjust the headphone volume for comfort, but only if required. For the combined Auditory-Tactile group, dual mono audio and vibrotactile stimuli were delivered to both the ears and hands in unison via Sennheiser headphones and a vibrating glove device [17]. Stimuli were applied to both hands simultaneously to control for increased dominant hand sensitivity or other variances of sensitivity that may have existed.

**Experiment Stimuli**

Waveforms were generated digitally using an open source wave editing software (Audacity) at a fundamental frequency of 160 Hz. The complex waveforms generated had infinite spectra and therefore could not be completely realised physically. However, it was possible to reshape waveforms through the transfer functions of the respective transducers. The phase and synchrony of the applied waveforms were kept constant by delivering the stimulus with the same onset time and with constant stimulus and interstimulus times (ISTs). Samples were arranged into five-second clips. Each clip consisted of a 2-second waveform, a one second IST, and a further 2-second waveform. The two waveforms varied in frequency from each other by ± 0.25, 0.5, 0.8, 1, 1.5, 2, 3, 4, 6, 8, 12 Hz. Each waveform clip was stored and presented to the participant three times during the experiment in counterbalanced order. Waveforms were outputted via a digital-analogue audio converter (Avid Fast Track C400) with a sampling frequency of 96 kHz and 24-bit resolution. The audio-only signal was routed through output channel one of the converter directly to the headphones. The same signal was also routed through output channel 2 and split to the left and right vibrating devices in parallel. Voltage measurements of RMS amplitude were captured and monitored at the input stage of the left-hand vibrating device.

**Waveform Types**

The auditory and vibrotactile stimuli applied during all experiment conditions were sine, saw, and square waveforms, with no aliasing for the square waveform, see Figure 1. As different musical instruments produce unique timbres, each instrument sounds quite different when presented at the same fundamental pitch. Therefore, the complex waveforms used in the study represented the different instrument tone qualities that a performer may be exposed to during a performance. The chosen waveforms displayed no harmonics (sine), odd harmonics only (saw), and odd and even harmonics (square) of the chosen fundamental. This allowed for the control of multidimensional aspects of waveform generation beyond frequency and amplitude while also considering timbre.

**Fundamental Frequency**

A fundamental frequency of 160 Hz was chosen for this experiment as it has been observed as having the lowest sub-threshold of perception in other experiments of this type. Furthermore, 160 Hz lies between the musical notes D3# and E3 (equal temperament scale), controlling for any advantage a musician may have had through experience. Waveform output levels from the test equipment to the vibrotactile gloves were pre-set to the following amplitudes: 0.03, 0.09, 0.31 Vrms for sine, saw, and square respectively. The ability to differentiate between waveform types in this way has been evaluated in other discrimination tasks [27]. Participants were also asked to verbally verify that the amplitudes of each of the tactile stimuli were perceptually equal during the initial setup period and trial stages of the experiment.

**Actuators and Arrangement**

The vibrating gloves were equipped with six independent audio-haptic voice-coil exciters, Figure 2. Similar voice-coil devices have been established as being suitable for transmitting the sonic characteristics of music in the form of vibration [17] [27] [28]. The voice-coil transducers have a diameter of 9mm and are designed to deliver vibrotactile output at frequencies the hand is most sensitive to. For each transducer a force factor of 2.4 Tm was outputted with a moving mass capability of 0.13g over a peak-to-peak maximum coil excursion range of 1.2mm [29].

**4 RESULTS**

For each participant, hit and false alarm data was transformed to calculate an independent observation of $d'$. 

![Figure 1: Waveforms applied to the transducers of the glove device.](image1.png)

**Figure 2: Voice-Coil locations on the glove device (palm region inset).**

![Figure 2: Voice-Coil locations on the glove device (palm region inset).](image2.png)
This value was then used to define an unbiased proportion of correct ‘Higher’ responses, $p(c)$ (Table 5.3 in Macmillan and Creelman, Detection theory: A user’s guide” [30]), and averaged across all participants. A logistic function of mean $p(c)_n$ was then applied to fit data to a psychometric function for each waveform type (equation 1), where $A$ and $B$ represent the parameters of a line, $f = \text{frequency}$, and $p = \text{the unbiased proportion of responses that } f \text{ was judged higher than } 160 \text{ Hz}$. Following this, JND was then calculated using equation 2.

$$A ( f - B ) = - \log \left[ \left( \frac{1 - p}{p} \right) \right] \quad \text{(equ.1)}$$

$$\text{JND}_{50} = \frac{B - 1}{A \left\{ \log \left[ \left( \frac{1 - 0.75}{0.75} \right) \right] \right\} } \quad \text{(equ.2)}$$

### TABLE 1: DESCRIPTIVE STATISTICS

<table>
<thead>
<tr>
<th>Waveform</th>
<th>Group</th>
<th>PSE</th>
<th>JND</th>
<th>r²</th>
<th>Mean</th>
<th>SD</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sine</td>
<td>Non-Musician</td>
<td>160.00</td>
<td>162.34</td>
<td>.86</td>
<td>.75</td>
<td>.09</td>
</tr>
<tr>
<td></td>
<td>Musician</td>
<td>159.97</td>
<td>162.07</td>
<td>.84</td>
<td>.77</td>
<td>.07</td>
</tr>
<tr>
<td>Saw</td>
<td>Non-Musician</td>
<td>160.00</td>
<td>162.24</td>
<td>.85</td>
<td>.76</td>
<td>.08</td>
</tr>
<tr>
<td></td>
<td>Musician</td>
<td>159.98</td>
<td>161.85</td>
<td>.80</td>
<td>.85</td>
<td>.09</td>
</tr>
<tr>
<td>Square</td>
<td>Non-Musician</td>
<td>160.00</td>
<td>162.04</td>
<td>.85</td>
<td>.80</td>
<td>.12</td>
</tr>
<tr>
<td></td>
<td>Musician</td>
<td>159.97</td>
<td>161.95</td>
<td>.83</td>
<td>.86</td>
<td>.12</td>
</tr>
<tr>
<td><strong>Auditory-Tactile</strong></td>
<td>****</td>
<td>****</td>
<td>****</td>
<td>****</td>
<td>****</td>
<td>****</td>
</tr>
<tr>
<td>Sine</td>
<td>Non-Musician</td>
<td>160.00</td>
<td>161.82</td>
<td>.79</td>
<td>.88</td>
<td>.07</td>
</tr>
<tr>
<td></td>
<td>Musician</td>
<td>159.98</td>
<td>161.75</td>
<td>.74</td>
<td>.94</td>
<td>.06</td>
</tr>
<tr>
<td>Saw</td>
<td>Non-Musician</td>
<td>160.00</td>
<td>161.75</td>
<td>.75</td>
<td>.92</td>
<td>.06</td>
</tr>
<tr>
<td></td>
<td>Musician</td>
<td>159.97</td>
<td>161.80</td>
<td>.80</td>
<td>.89</td>
<td>.08</td>
</tr>
<tr>
<td>Square</td>
<td>Non-Musician</td>
<td>160.00</td>
<td>161.73</td>
<td>.76</td>
<td>.94</td>
<td>.04</td>
</tr>
<tr>
<td></td>
<td>Musician</td>
<td>159.97</td>
<td>161.73</td>
<td>.76</td>
<td>.94</td>
<td>.04</td>
</tr>
</tbody>
</table>

### TABLE 2: TWO-WAY BETWEEN GROUPS ANOVA

<table>
<thead>
<tr>
<th>Interaction Effect</th>
<th>Type III Sum of Squares</th>
<th>df</th>
<th>Mean Square</th>
<th>F</th>
<th>Sig.</th>
<th>Partial Eta²</th>
</tr>
</thead>
<tbody>
<tr>
<td>Grouping &amp; Musicianship</td>
<td>.001</td>
<td>1</td>
<td>&lt;.001</td>
<td>.013</td>
<td>.91</td>
<td>&lt;.001</td>
</tr>
<tr>
<td><strong>Main Effect</strong></td>
<td>****</td>
<td>****</td>
<td>****</td>
<td>****</td>
<td>****</td>
<td>****</td>
</tr>
<tr>
<td>Grouping</td>
<td>.205</td>
<td>1</td>
<td>.21</td>
<td>26.68</td>
<td>&lt;.001</td>
<td>.25</td>
</tr>
<tr>
<td>Waveform</td>
<td>.025</td>
<td>2</td>
<td>.01</td>
<td>1.56</td>
<td>.22</td>
<td>.04</td>
</tr>
<tr>
<td>Musicianship</td>
<td>.078</td>
<td>1</td>
<td>.08</td>
<td>9.954</td>
<td>.002</td>
<td>.11</td>
</tr>
</tbody>
</table>

From the results presented, it can be observed that the detection of auditory changes in the order of ± 12 Hz around a fundamental frequency of 160 Hz can be facilitated by the simultaneous crossmodal presentation of vibrotactile stimuli. When auditory feedback was combined with vibrotactile feedback there was seen to be a statistically significant improvement in the Audio-Tactile group’s ability to discriminate between auditory frequency variations above that of levels when auditory stimulation was presented alone. A two-way between-group analysis of variance was conducted to explore the impact of test grouping and musicianship on the unbiased proportion of correct ‘Higher’ responses (table 2). The interaction between the independent variables of grouping and musicianship was found to have no significant effect.

The main effect for grouping reached statistical significance. Additionally, there was also found to be a significant increase in frequency discrimination within both groups for musicians. These findings support suggestions that there is a close relationship between auditory and somatosensory stimulation in the auditory cortex of the brain. A relationship that has been observed in FMRI studies that capture the mapping of audio-tactile co-activation in the auditory belt of the brain [6].

### Figure 3: Psychometric functions for sine, saw, and square waveforms.

### Figure 4: Box Plots representing median $p(c)_{50}$ across all waveforms and musicianship (outliers indicated by circles).

#### 5 DISCUSSION

Although the main effect of waveform was not found to be significant, the sinewave stimuli presented with a much more distinct curve between groups than for both of the complex waveforms. This indicated that in the application of extra-auditory vibrotactile feedback in pitch detection exercises, the complexity of the waveform has some influence upon the perception of pitch. That is to say, the effect was less noticeable for more complex waveforms. However, this does not diminish the potential application of complex waveforms in vibrotactile feedback, but suggests that in real-world applications a balance between simple and complex waveforms must be explored.

Similarly, interesting observations were made for expected values of JND, as the JND of the tactile system is much broader than that of the auditory. For example, the expected tactile only JND of a 150 Hz sinusoidal stimulus with the amplitude held constant has been measured as ± 18% (27 Hz) of the fundamental [31], equating to 28.8 Hz at 160 Hz. In addition, in an auditory only JND experi-
ment there would be expected a 3 Hz variation in JND for sine wave and 1 Hz for complex waveforms below 500 Hz [32]. As can be seen in Table 1, the JND results for the Audio-Only group presented with an average of 2.33 Hz for sine, 2.22 Hz for saw, and 2.06 Hz for square waveforms. In the combined Audio-Tactile group, for the sine waveforms, the JND was measured at 1.83 Hz and the observed JND for the complex waveforms measured as 1.89 Hz and 1.78 for saw and square respectively. This indicated that the JND for all waveforms was perceived relatively equal, with only a small improvement when vibrotactile information was included. Although there is a relatively broad JND for the tactile system, when combined with auditory stimuli, there appeared to be some small practical effect upon this group’s average JND values for all waveforms.

Applications in Music.

As discussed earlier, acoustic instruments allow musicians to create sound through explicit gestures that are specific to the sound generator that the particular instrument employs. Conventionally, they are designed for single users, are single-sound oriented or sound specific in their design, and the context in which they are used is largely determined by the user. As the input and the output of an instrument is physically inseparable, an explicit dialogue is formed between the musician and the instrument, one that is established through extended practice and performance. This relationship is used to facilitate the user when they apply interactions learned from one instrument to others of similar design.

For many new DMI’s, these relationships are not so apparent. New interfaces for musical expression are becoming evermore multi-modal and embedded, allowing musicians to interact with digital sound generating modules in a multitude of novel and innovative ways. In many instances, haptically enabled DMIs allow for a more natural or familiar interaction to take place between computers and musicians, bridging the physical-digital divide with an interaction/feedback pattern that is recognizable to the user. Furthermore, instead of creating computer interfaces for musicians, DMI designers now have the potential to provide musician interfaces for computers. The nature of the interaction is changing beyond traditional concepts of a musical interaction, yet there is still the possibility to stimulate the user in an evocative and familiar way. If future DMI designers continue to neglect the potential of vibrotactile feedback to tap into the deeper philosophical potential of musical haptics, a hypothetical distance between the user and the systems in use will be created and increase the disconnect felt between the user and the digital world. Accordingly, this physical-digital divide will continue to present interface designers with issues beyond basic interaction metaphors and it is therefore suggested that DMIs should be developed to crossmodally stimulate a user when pitch specific exercises are being executed.

Vibration perception in musical interactions can make use of both active and passive feedback during acoustic performances. While perception is usually the main focus when designing a haptically enabled DMI interface, we suggest that it should also be complemented with a relatable passive feedback element to create a more recognisable multimodal interaction. Furthermore, in musical performances, the application of audio related tactile feedback bridges passive with active feedback as the instrumentalist in motion is actively interacting with the source of the passive feedback. In new musical devices, it is in the decoupling of the gestural controller from sound generator that the role of passive and active feedback becomes separated. Information about process is often lost and the relationship with the gesture captured is simulated or loses transparency [33] [34]. It is therefore suggested that this may be avoided with the inclusion of vibrotactile feedback in DMI design.

6 CONCLUSIONS

In the experiment presented, the role of extra-auditory vibrotactile feedback was investigated. It was initially hypothesised that vibrotactile feedback may be applied to pitch selection exercises to further enhance DMI feedback practices. Through the application of extra-auditory feedback in a 2AFC task, the participants displayed an increased awareness of pitch variation. In consideration of these findings and their potential application to haptic interactions, it is recommended that the adoption of a combined psychophysical model is required to reinforce the role of somatosensory integration in frequency discrimination tasks that are carried out on digital musical instruments. This will potentially allow researchers to design multisensory interfaces that are transparent and intuitive for musicians to use during performances.

REFERENCES


