Sonification of virtual and real surface tapping: evaluation of behavior changes, surface perception and emotional indices

Ana Tajadura-Jiménez[1], Enrico Furfaro[1], Nadia Berthouze[1] and Frédéric Bevilacqua[2]

[1] UCL Interaction Center, University College London, London, UK

[2] STMS Lab, IRCAM-CNRS-UPMC, Paris, France

Word count (including abstract, main body & reference text, but no caption text or figures and tables.): 5639

To whom correspondence should be addressed:

Dr Ana Tajadura-Jiménez, UCL Interaction Centre (UCLIC), Malet Place Engineering Building, 8th floor; University College London; Gower Street, London WC1E 6BT, UK. Tel. +44(0) 20 7679 0683 (30683). Fax. +44(0)1784434347, E-mail: a.tajadura@ucl.ac.uk

Abstract

The audio-feedback resulting from object interaction provides information about the material of the surface and about one's own motor behavior. With the current developments in interactive sonification, it is now possible to digitally change this audio-feedback, and thus, the use of interactive sonification becomes a compelling approach to shape tactile surface interactions. Here, we present a prototype for a sonic interactive surface, capable of delivering surface tapping sounds in real-time, when triggered by the users' taps on a real surface or on an imagined, "virtual" surface. In this system, the delivered audio-feedback can be varied so that the heard tapping sounds correspond to different applied strength during tapping. Here, we also propose a multi-dimensional measurement approach to evaluate user experiences of multi-modal interactive systems. We evaluated our system by looking at the effect of the altered tapping sounds on emotional action-related responses, users' way of interacting with the surface, and perceived surface hardness. Results show the influence of the sonification of tapping at all levels: emotional, behavioral and perceptual. These results have implications in the design of interactive sonification displays and tangible auditory interfaces aiming to change perceived and subsequent motor behaviour, as well as perceived material properties.

Keywords: Sonification; Multimodal interfaces; Interaction styles; Surface interaction; Emotion; Evaluation Methods

Our interactions with objects are characterized by the sensory feedback that accompanies these interactions. For instance when we touch the surface of a table, we may see the surface and our hand moving on top of it, at the same time we get tactile cues that change as we move our hand, and we may also hear the interaction sound produced by our hand rubbing the surface. In fact, the perception of materials is known to be multisensory, with touch, vision, and audition all contributing to it and interacting with each other[1]. These sensory cues inform of properties of the objects, such as texture, shape or hardness, and they also inform of our interaction behavior, such as the speed, the amplitude and fluidity of our movements, as well as the strength we are putting into them.

Current developments in multi-modal interactive systems allow for digitally changing the sensory cues resulting from our interaction with objects. This opportunity opens new avenues in the use and design of both physical and virtual objects. For example, the altering of the sound an object makes when we scratch its surface with eyes closed may lead us to perceive the object as rougher and consequently to increase the strength of our scratching behavior. Our emotions may also be altered in response to the altered audio-feedback and to the changes in proprioceptive and tactile feedback. In fact, various studies have shown a tight link between body movement and emotions. For instance, body movement biases one’s mood towards the mood that the movement expresses[2,3].

Multi-modal interactive systems with the potential to change users’ perception of properties of the objects, motor behavior and emotional state have applicability for the design of technology in numerous contexts. As interaction with objects is increasingly mediated through their digital representation, audio-feedback can be used to complement the limited amount of haptic feedback available to understand the object properties and facilitate its virtual or remote handling. For instance, in the context of online shopping, the perceived properties of materials and emotional responses are leading decisions factors[4]. Another application example is touch-less surgery, where extreme precision in applied strength is required, and it is important that information about the material properties of the objects manipulated is fully provided, and even enhanced, to facilitate the risky surgery process.

Multi-modal interactive systems might be also used in the contexts of fun-promoting and health-promoting applications, such as videogames, physical or mental rehabilitation apps, where specific ways of performing movements are fundamental to reach specific objectives. Providing wider sensorial experiences may impact on cognitive processes, may help to reduce the overall mental effort required to operate the system, and may induce more engaging and more intense emotional experiences. Evidences from various studies have shown that affective touch and movement behaviour profiles do exist[2]. By using mechanisms to alter touch and motor behavior, game designers are provided with ways to modulate or enhance the player's emotional experience.

In the context of physical therapy, inducing motor behaviour changes in a self-controlled way may reduce the danger of over stress on limbs in the absence of physiotherapists. It may also increase the perceived self-efficacy by making one feel stronger, faster or more pleased, which will eventually impact on motivation and adherence to therapy[5]. Nevertheless, research needs to be conducted in order to set the bases for the design of these systems. In particular, an exploration is needed into the various mapping strategies for creating coherency between action and reaction, and into ways of evaluating the performance and usability of the designed systems.

This article addresses the use of interactive sonification[6] as a compelling approach to shape tactile surface interactions. Sonification of actions, consisting of the mapping of gestures and actions into sound, is rather a new approach in the design of multi-modal interactive systems, as compared to visualization, but it is a powerful one. In particular, in interactive systems, audio feedback has generally been used to notify the user of success or failure of an action, instead of sonifying the action in itself. As a proof-of-concept, we present a prototype that we designed that allows for the sonification of surface tapping. This system is capable of delivering surface tapping sounds in real-time when triggered by the users' taps on a real or on an imagined, "virtual", surface (i.e., when tapping in the air). Having real and virtual surface types allows exploring the effects of audio-feedback when tactile cues informing of the tapped surface and of applied strength are either present or absent.

We, then, describe the design and procedure followed in the system evaluation. Multi-modal interactive systems in general, and interactive sonification systems in particular, are often poorly evaluated. In this article we argue for adopting a multi-dimensional measurement approach to evaluate user experiences of multi-modal interactive systems. This approach may combine self-reporting, physiological measurements and objective behavioral data. Many systems might be evaluated using only one of these measures, but the combination of measures brings us closer to an understanding of the potential effects of the system, which will inform its design. To evaluate our system, we quantified changes in perception of surface hardness, tapping behaviour and emotional action-related responses. Our results show the power of sonification to induce changes at all these levels.

Audio-feedback during object interaction

When people touch or tap on a surface, they can often hear the resulting interaction sounds[7]. Different physical features of the material of the surface will result in different auditory cues; for instance, tapping on a soft woollen surface will produce different sounds than tapping on a hard wooden surface. Different modes of touching the surface will also result in different auditory cues; for instance, tapping soft on a surface will produce *weaker* sounds than tapping hard on the same surface. But to what extent do we make use of this information available during surface interaction sounds?

Several recent studies have shown that changing the audio-feedback resulting from object surface interaction may lead to changes in the perceived material properties of the objects, both in the case of natural surfaces[8] and of virtual haptic surfaces[9]. Other studies have shown that providing altered audio-feedback may also lead to a change in our way of interacting with the objects. For instance, hearing the expected contact sound on the onset of a reaching-to-grasp movement towards an object (i.e., hearing the sound that touching that object would produce) can speed the movement as compared to when hearing an unexpected contact sound (i.e., the sound of an object with different material)[10].

Importantly, altering the audio-feedback during object interaction may change motor behavior because the feedback informs of the motor behavior itself, as well as of properties of one's own body. For example, sonification of boat motion improves movement execution of elite rowers, as it provides information about small variations and deviations in rowers' movements[11]. Sonification of tapping actions can actually change the perceived length of one's arm tapping on a surface, as the tapping sounds inform of the location and dimensions of the arm[12]. Introducing a delay in the footsteps sounds produced when walking results in changes in gait-period and walking speed[13]. Moreover, altering footstep sounds to represent walking on different ground surfaces seems to influence the walking style when people intend to walk with specific emotion-related styles[14].

In this article, we advance these studies by focusing on altering the audio-feedback related to the level of applied strength when tapping on a surface, rather than focusing on the feedback related to specific materials, and without asking any specific behaviour style. We designed a prototype that allows for the sonification of surface tapping. This system is capable of delivering, via headphones, surface tapping sounds in real-time when triggered by the users' taps on a real surface (a table top) or on an imagined, "virtual", surface (when tapping in the air). This system thus allows exploring the effects of audio-feedback when tactile cues informing of the tapped surface and of applied strength are either present (real surface) or absent (virtual surface). Although the action of tapping on a virtual, non-existing, surface is physically not possible, it is still possible to imagine this surface and to perform a similar action; this action is facilitated by the audio-feedback that indicates when the virtual surface is hit. Virtual objects are becoming part of our everyday environments and it is therefore important to understand how they are handled. Figure 1A displays examples of a person tapping on these two surface types.

[Insert Figure 1 about here]

By means of interactive sonification of surface tapping actions, we aimed to explore how sounds produced when tapping on a surface actually (1) inform of the physical feature of hardness of the surface material; (2) inform of the applied strength when tapping; (3) inform of the user's ability to tap, which may impact on their own emotional state; and (4) change user's own tapping behaviour, as they will try to adjust their tapping actions in response to the audio-feedback, an effect often referred to as auditory-action loop.

Sonification of virtual and real surface tapping: System overview

Sonification of surface tapping for the system we designed is achieved by having the tapping action triggering the presentation of pre-recorded tapping sounds in *real-time* (the mean delay introduced by the system is 10.7 +/- 1.8 ms; the maximum delay is 14 ms). The tapping action is detected by registering the signals captured by a piezoelectric transducer, attached to the "real" surface, and by an accelerometer, attached to the middle finger of the users' dominant hand. An overview of the connections of the prototype physical components is displayed in Figure 1B.

We use a motor-to-audio translation algorithm that triggers a feedback sound every time a "real" or "virtual" tap is detected. For the detection of surface taps a threshold is set as follows. For the "real" surface condition the threshold is based on the absolute value of the peak amplitude of the piezo input signal, being specifically calibrated according to the piezo sensitivity to detect surface taps. For the "virtual" surface condition, since the hand is kept in the air, the sound is triggered using the accelerometer signal. The trigger depends on a threshold set on the derivative of the accelerometer signal. The pre-recorded feedback sound is the sound produced by a person tapping on a surface. Across conditions the feedback can be varied so that the heard tapping sounds correspond to different applied strength during tapping.

The system allows recording the piezo and accelerometer input signals, as well as the generated audio-feedback, which can be used to analyze user's tapping behavior (i.e., maximum acceleration and frequency of users' tapping movements).

We calibrated the system according to the accelerometer and piezo input ranges, and to remove the background noise in the piezo signal. A sensor attached to the user's wrist (non-dominant hand), measures the galvanic skin response (GSR) of the user. GSR is a sensitive and valid real-time measure for emotional arousal in response to external stimuli[15].

System evaluation

In order to evaluate our system we recorded three sounds in an anechoic chamber, which allowed reducing background noise. The duration of the sounds was 190 ms, and the sampling rate used 44.1 kHz. The sounds were of a person tapping with the palm of the hand on a cardboard box applying three different levels of strength. We chose the sound of tapping on a cardboard box given the rather clear difference in sounds resulting from different levels of applied tapping strength. We refer to these three versions of the sounds as "weak", "medium" and "strong" tapping sounds. The sounds were normalized using Audacity software so that there was an 8 dB difference between "weak" and "medium" sounds, and between "medium" and "strong" sounds.

We asked 23 participants (5 male, 18 female; 19-35 years old; mean age 23.2) to take part in the evaluation. All participants reported having normal hearing and normal tactile perception. They were blindfolded, except two participants that preferred to keep their eyes closed. They were required to tap onto the two types of surfaces, real and virtual, while receiving audio-feedback in response to their tapping actions. We followed a within-subjects design, with all participants being exposed to all sound conditions, presented in randomized order. In particular, each participant completed six tapping blocks differing in the type of tapped surface (surface type:

real or virtual) and the level of strength conveyed by the tapping sounds presented as feedback (sound strength level: weak, medium or strong).

Figure 1C displays the timeline for each experimental block. Each block lasted for 80 s. Participants were asked to tap with their dominant hand on the real or the virtual surface for the whole duration of the block. They were required to keep their rhythm constant and to produce one tap approximately every second. We specifically asked participants to maintain the same tapping style across the experimental blocks. During the first and last 10 s of the block, which we called baseline1 and baseline2, participants only heard pink noise. For the remaining time of the block, which we called feedback phase, apart from pink noise, participants were presented with real-time audio-feedback in response to their taps. GSR was recorded during the whole duration of the block. At the end of each block, participants filled in a questionnaire that allowed us to assess their subjective experience during the block.

Before each experiment, we made sure that all input signals (piezo, accelerometer, GSR) were detected. In particular, we tested the GSR recordings by looking at signal changes in response to the participant taking a deep breath and in response to a sudden noise.

Sounds, MAX/MSP patches, questionnaire and data collected are available at: https://www.ucl.ac.uk/uclic/research/project-pages/hearing-body/Furfaro_IEEEMultimedia.

Multi-dimensional measurement of user experience

One of the aims of this article is to demonstrate the use of a multi-dimensional measurement approach to evaluate User Experience (UX). We can broadly classify the measurement dimensions into three categories attending to whether they look at alterations in perceptual aspects of the experience (e.g., surface perception), alterations in behavior or alterations in emotional experience. Furthermore, the latter can be quantified by looking at subjective, behavioral or physiological emotion-related changes. Each of these dimensions tackles a different UX aspect, which may or may not correlate with each other. Hence, when possible, it is strongly recommended that the UX evaluation, when interacting with multi-modal interactive systems, combines measures taken at the three different levels. This multi-dimensional measurement approach does not always imply increasing the UX complexity and might help to bring us closer to an understanding of the potential effects of a given system, which will in turn inform its design. Here we present a practical example of such multi-dimensional measurement approach to demonstrate its feasibility and the richness of the information on UX it provides.

Our hypothesis was that our system, by altering the audio-feedback cues that inform of the applied strength when tapping on the surface, induces changes on perceived applied strength when tapping, perceived one's own ability to tap and other emotional responses to the tapping task; changes on tapping behaviour; and changes on perceived surface hardness. Hence, for the UX evaluation of our system we looked at alterations at all these different levels.

We looked at changes in emotional action-related responses, by quantifying subjective and physiological emotion-related changes. For this purpose, in the questionnaire we included several scales. First, we used 7-point Likert scales assessing the perceived physical strength, ability to complete the task, and aggressiveness felt when tapping on the surface. Second, we also quantified the subjective mental effort, by asking participants to indicate the stress felt while tapping, using a vertical analog scale[16]. Third, we quantified emotional valence, dominance, and arousal felt by participants by using the three 9-item graphic scales of the self-assessment manikin[17]. Arousal was also quantified by looking at the physiological changes recorded by the GSR biosensor. We looked at alterations in the way of interacting with the surface by quantifying the changes in the movement dynamics. For this purpose, we used the logged accelerometer and piezo data. Finally, we assessed the perceived surface physical quality of hardness with a 7-point Likert scale.

Each of these dimensions tackles a different UX aspect, but these aspects may correlate with each other. For instance, induced changes on perceived applied strength when tapping, or the changes on perceived ability to tap, may be accompanied by corresponding changes in tapping behaviour. Therefore, in order to understand whether the different dimensions of experience are linked, we performed correlation analyses between the different measures.

All data collected were statistically analyzed with Statistical Package for the Social Sciences (SPSS) 21.0. Shaphiro-Wilk tests assessed normality of data distributions. Parametric (analysis of variance – ANOVA - and t-tests), and non-parametric (Friedman and Wilcoxon) tests were used, respectively, with normal and non-normal data.

Sonification of surface tapping changes emotion and surface perception

Figure 2A shows the mean self-reported perceived aggressiveness, perceived physical strength, ability to complete the tapping task and perceived surface hardness; Figure 2B shows the self-reported valence, arousal and dominance; Figure 2C shows the perceived effort while tapping; and Figure 2D shows the GSR change scores, all according to the level of strength of the tapping sounds presented as feedback (weak, medium and strong) and the type of tapped surface (real or virtual). We first report the effects due to the sound strength level, and then the effects due to the surface type.

[Insert Figure 2 about here]

Changing the strength level conveyed by the sound, contrary to our expectations, did not alter the perceived applied strength or the aggressiveness felt when tapping (all *ps* > 0.05). However, changes in perceived applied strength and aggressiveness did correlate with changes in other dimensions, such as behavioral changes, perceived surface hardness and several emotional dimensions (see the section "Correlation between measures"). In addition, results show that, when tapping on a real surface, participants felt less able to tap and less pleasant in the case of the low intensity sound. In particular participants felt less capable to tap for the *weak* than for the *medium* condition ($z = -2.12$, $p < 0.05$), and found that the experience of tapping was less pleasant for the *weak* than for the *strong* sound condition ($z = -2.31$, $p < 0.05$). We also show that, when tapping on a virtual surface, participants felt more physiologically aroused when the sound informed of low level of tapping strength. In particular participants GSR was higher for the *weak* than for the *strong* ($z = -4.01$, $p < 0.001$) or *medium* sound condition ($z = -4.05$, $p < 0.001$). This result highlights that audio-feedback related to tapping strength informs users of their performance and that emotional experience is affected by the congruence between tapping sounds and tapping actions. Audio-motor incongruences lead to unpleasant and arousing experiences. When the audio-feedback does not match the expectations of people, as it happens in the case of the *weak* sound, participants feel more aroused or stressed, and less able to tap. In the case of the *weak* sound, the incongruence between the applied strength and the sound heard as output becomes more evident, as the sound heard was produced by applying very little strength.

Sound did not change even if participants explored different movement strategies in this condition(see section below on “sonification of surface tapping changes behaviour” and Figure 3), which further contributed to participants realizing the incongruence. The coherence between an action and its auditory response is known as one of the principles of altering interaction sounds to successfully convey information and modulate actions in an intuitive manner (see work on “Blended sonification”[18]).

Figure 2A reveals a significant correlation between the sound strength level and the mean perceived surface hardness. We found that when no tactile cues are available (i.e., virtual surface), participants make use of the audio-feedback to decide on the hardness of the material being tapped. In particular, participants seem to match the level of strength applied when tapping, as conveyed by sound, with the level of hardness of the surface. Participants perceived the tapped surface as being softer for the *weak* than for the *strong* ($z = -2.34$, $p < 0.05$) and the *medium* ($z = -2.21$, $p < 0.05$) sound conditions. No such results were found for the real surface condition, which provides additional tactile cues about the surface. Differences between conditions in which a surface is explored by sound and finger touch, as opposed to when no finger touch is available, have been previously reported. For instance, sound feedback seems more informative of the roughness of the texture of a surface when the surface is inspected with a rigid probe than when inspected by the fingers[1].

When looking at the effects due to the surface type, we found, as expected, differences between the real and virtual surface conditions at all measured levels. The real surface was perceived as harder and caused feelings of greater strength, of larger ability to tap, and of being less stressed when tapping. First, for all sound conditions, participants perceived the tapped surface as being harder when tapping on the real than in the virtual surface (*strong*: $z = -3.48$, $p = 0.001$; *medium*: $z = -3.31$, $p = 0.001$; *weak*: $z = -3.81$, $p < 0.001$), as it can be seen in Figure 2A. Second, in the same figure it can be observed that our participants felt they applied more strength when tapping on the real than on the virtual surface, at least for the *medium* sound condition ($z = -1.98$, $p < 0.05$). Third, for most sound conditions, participants felt more able to tap (*medium* sound: $z = -2.98$, $p < 0.005$; *weak* sound: $z = -2.23$, $p < 0.05$) and less aroused when tapping on a real rather than virtual surface. The arousal-related results were confirmed by looking at the self-reported arousal (*strong* sound: $z = -2.28$, $p < 0.05$; *weak* sound: $z = -2.17$, $p < 0.05$), and at the physiological GSR recordings (*weak* sound: $z = -4.17$, $p < 0.001$). The increase in arousal in the virtual surface may come from the unnaturalness of such an interaction, compared to the more common interaction with a real surface. In addition, the observed differences between the effects of tapping on real and virtual surfaces might relate to the fact that during the real surface conditions there were also tactile cues present, additionally to auditory and proprioceptive cues. One should bear in mind that the perception of materials, and our perception in general, is known to be multisensory, with all sensory modalities contributing to it and interacting with each other[1]. In this case, participants were blindfolded, and thus, visual cues were not available, but auditory, proprioceptive and, in the case of the real surface, tactile cues, contributed all to surface perception.

Sonification of surface tapping changes behavior

We show that by presenting real-time audio-feedback regarding tapping strength, we can actually change the tapping behavior when tapping in both real and virtual surfaces. From the accelerometer values, we can estimate parameters that relate to the movement dynamics. First, accelerometer data reveal how hard the participants hit the real surface or stopped their motion in the case of the virtual surface. We can also quantify the inter-tapping intervals. These different measures are reported for baseline1, baseline2 and feedback phase, and are displayed in Figures 3A and 3B.

[Insert Figure 3 about here]

Separate analyses for each phase showed an effect of surface for baseline1 ($F(1, 22) = 9.89, p = .005$), and no significant effects for baseline2 (all *p*s > 0.05). Importantly, for the feedback phase, we found that according to the audio-feedback received, participants changed their own motor behavior. In particular, the acceleration maxima were significantly affected when the sound suggested that a low strength level had been applied when tapping, as compared to when it suggested a high strength level ($p < 0.05$).

Figures 3C and 3D show how the acceleration changes from baseline1 across time. Looking at these figures allows better interpreting the changes in participants' behavior. It can be observed that the *medium,* and specially the *weak,* conditions are accompanied by much bigger changes in acceleration, as well as by much bigger differences between participants, than the *strong* sound condition. This might be interpreted as behaviour being affected by the participants' expectations of the motion-sound interaction. In our system, the congruency between tapping sounds and tapping actions is highly non-linear: while the *strong*, and to some extent also the *medium*, sound conditions might appear as possibly congruent, the *weak sound* condition is highly incongruent with participants' actions. This incongruency results in participants exploring different movement strategies (e.g., trying to stop the hand or to put more strength into their taps) in an attempt to compensate for the *weak* audio-feedback.

Analyzing the tapping behavior in terms of differences between phases allowed investigating the overall effect of audio-feedback in tapping behavior. This effect was significant both in terms of acceleration ($F(2, 44) = 10.72$, $p < 0.001$), as displayed in Figure 3A, and in terms of tapping frequency ($F(2, 44) = 10.24$, $p < 0.001$), as displayed in Figure 3B. Comparing the effects during the feedback phase with those during the first period of tapping (baseline1), where participants did not receive audio-feedback, showed that overall, introducing audio-feedback, regardless of the level of strength conveyed, speeded participants' movements ($p = 0.001$) and decreased the acceleration ($p < 0.001$). Interestingly, these effects seem to persist after 60 seconds of audio-feedback, even when audio-feedback is not present anymore. This is confirmed by the non-significant differences between the feedback and baseline2 phases (all *ps* > 0.05), which might indicate some adaptation or persistence of the audio-feedback effect.

Finally, there was also a significant interaction in terms of movement acceleration between surface type and phase ($F(2, 44) = 9.02$, $p = 0.001$), showing that while for the real surface condition there were differences between baseline1 and feedback phase ($p < 0.01$) and baseline 2 ($p < 0.05$), these differences were not observed for the virtual surface condition (all *ps* > 0.05).

Other studies have shown similar auditory-action loops that can result in changes in movement execution, for instance when rowing[11] or walking[13,14]. However, here we show that by presenting real-time audio-feedback related to the tapping strength, we can indeed change the tapping behavior. Changes occur even in a virtual environment, where the surface on which tapping is performed is simulated.

Correlation between measures

Finally, in order to further understand how the different dimensions of experience are linked, we performed correlation analyses in which we looked at the changes across measures due to the sound strength level, from *strong* to *weak*. These correlations are presented in Table 1. We highlight the fact that changes in behavior (acceleration patterns) were accompanied by changes in perceived applied strength, in both real ($r = 0.55$, $p < 0.01$) and virtual surfaces ($r = -0.42$, $p < 0.05$), although the direction of change varied with surface type. Behavioral changes were also accompanied by changes in perceived surface hardness for the real surface ($r = 0.62$, $p < 0.005$). In addition, perceived surface hardness correlated with other emotional measures, such as dominance ($r = 0.58$, $p < 0.005$) and aggressiveness ($r = -0.44$, $p < 0.05$) for the virtual surface condition, and perceived physical strength for the real surface condition ($r = 0.69$, $p < 0.001$). Self-reported feelings of control (i.e., emotional dominance) correlated with physiological arousal for the virtual surface (GSR; $r = -0.43$, $p < 0.05$). One can also see how the different self-reported measures of emotion correlated at different levels.

[Insert Table 1 about here]

Conclusions

In this article we addressed the use of interactive sonification as a powerful tool to shape tactile surface interactions, as well as the use of a multi-dimensional measurement approach to evaluate user experiences of multi-modal interactive systems. We presented a prototype that allows for the sonification of surface tapping. This prototype is able of detecting user taps on a surface and of triggering in real-time the presentation of pre-recorded tapping sounds. In our system, it is possible to vary the delivered audio-feedback so that the heard tapping sounds correspond to different applied strength during tapping. This system can be used when tapping both on real surfaces (e.g., a table) and on imagined, “virtual” surfaces (i.e., when tapping in the air). We found that when the level of tapping strength conveyed by the audio-feedback was varied, participants experienced behavioral, emotional and perceptual changes due to the audio-feedback. These changes correlated with the changes in applied strength reported by participants.

More research is of course necessary to apply such findings to the design of technology. In order to further explain the differences between the real and virtual surfaces it would be in fact interesting to perform a more detailed analysis of the behaviour changes, for instance looking at the acceleration before the shock on the table occurred, and also looking at the envelope of the movement signals to understand the behaviour changes both when moving the hand upwards and downwards. It would also be interesting to test the system when using sounds that can induce larger emotional responses or result in a more aggressive behaviour. The present results are however very promising, as they open new avenues for research aiming to change movement behaviour, emotional state and material perception, in both real and virtual environments. Future research should further explore these effects and their applications, by combining both quantitative and qualitative multi-dimensional measurement methods to better understand the effects and possibilities these mechanisms provide.

These results may be applied in the design of interactive sonification displays and tangible auditory interfaces aiming to change perceived and subsequent motor behaviour, as well as perceived material properties.

References

1. R.L. Klatzky, and S.J. Lederman, “Multisensory texture perception”, in J. Kaiser, M.J. Naumer (Eds), Multisensory object perception in the primate brain, New York: Springer, 2010, pp. 211-230.
2. Y. Gao, N. Bianchi-Berthouze, and H. Meng, “What does touch tell us about emotions in touchscreen-based gameplay?”, in ACM Transactions on Computer - Human Interaction 19(4), pp. 31, 2012.
3. N. Bianchi-Berthouze, “Understanding the role of body movement in player engagement”. Human Computer Interaction 28(1), 2012, pp. 42-75.
4. J. Peck and J. Wiggins, “It just feels good: Customers’ affective response to touch and its influence on persuasion”. Journal of Marketing, 70(4), 2006, pp. 56-69.
5. M.S.H. Aung, B. Romera-Paredes, A. Singh, S. Lim, N. Kanakam, N., A.CdeC. Williams, and N. Bianchi-Berthouze, “Getting rid of pain-related behaviour to improve social and self perception: a technology-based perspective”, in Proc. IEEE International Workshop on Image and Audio Analysis for Multimedia Interactive Services (WIAMIS) 2013.
6. T. Hermann, A. Hunt, and J.G. Neuhoff, The sonification handbook. Berlin: Logos Publishing House, 2011.
7. D. Rocchesso, and F. Fontana (Eds.), The sounding object. Mondo estremo, 2003.
8. D. Merrill, and H. Raffle, “The sound of touch”, in Proc. ACM Conference on Human Factors in Computing Systems 2007, ACM Press, 2007, pp. 2807-2812.
9. M.R. McGee, P.D. Gray, and S.A. Brewster, “Mixed feelings: Multimodal perception of virtual roughness”, in Proc. Eurohaptics 2002, pp. 47-52.

10. U. Castiello, B.L. Giordano, C. Begliomini, C. Ansuini, and M. Grassi, “When ears drive hands: the influence of contact sound on reaching to grasp”. PLoS ONE, 5, 2010.
11. N. Schaffert, K. Mattes, and A.O. Effenberg, “Listen to the boat motion: acoustic information for elite rowers”, in Proc. Interaction Sonification workshop (ISon), Stockholm, Sweden, 2010.
12. A. Tajadura-Jiménez, A. Väljamäe, I. Toshima, T. Kimura, M. Tsakiris and N. Kitagawa, “Action sounds recalibrate perceived tactile distance”. Current Biology, 22(13), 2012, pp. R516-R517.
13. F. Menzera, A. Brooksab, P. Haljea, C. Fallera, M. Vetterlia, and O. Blanke, “Feeling in control of your footsteps: Conscious gait monitoring and the auditory consequences of footsteps”. Cognitive Neuroscience, 1(3), 2010, pp. 184-192.
14. R. Bresin, A. de Witt, S. Papetti, M. Civolani, and F. Fontana, “Expressive sonification of footstep sounds”, in Proc. Interaction Sonification workshop (ISon), Stockholm, Sweden, 2010.
15. W. Boucsein, Electrodermal activity. New York: Plenum press, 1992.
16. J. Sauro and J.S. Dumas, “Comparison of three one-question, post-task usability questionnaires”, in Proc. ACM Conference on Human Factors in Computing Systems 2009, ACM Press, 2009, pp. 1599-1608.
17. M. Bradley and P.J. Lang, “Measuring emotion: the Self-Assessment Manikin and the Semantic Differential”. Journal of Behavior Therapy and Experimental Psychiatry, 25, 1994, pp. 49-59.
18. R. Tuennermann, J. Hammerschmidt, and T. Hermann, “Blended Sonification: Sonification for casual information interaction”, in Proc. International Conference on Auditory Display 2013, Lodz, Poland, 2013.

Figure Legends

Figure 1. (A) A person using the prototype we designed for tapping on the "real" surface (top panel) and an example of the hand movement when tapping on the "virtual" surface (bottom panel). (B) Overview of the connections of the prototype physical components. The tapping actions are detected by using a piezoelectric transducer (Schaller Oyster 723 Piezo transducer Pickup), attached to the real surface, and an accelerometer (Triple Axis Accelerometer Breakout MMA8452QA), attached to the user's middle finger. The piezo is connected to an external soundcard and the accelerometer is connected to an Arduino Uno microcontroller board. Both connect through USB ports to a computer running the real-time synthesis environment MAX/MSP. Every time a "real" or "virtual" tap is detected, a pre-recorded feedback sound is played through headphones. The feedback can be varied to present tapping sounds that correspond to different applied strength during tapping. Piezo and accelerometer input signals, as well as the generated audio-feedback, are recorded and used to analyze user's tapping behavior. The audio-feedback is played back using closed headphones with high passive ambient noise attenuation (Sennheiser HDA 200). The use of these headphones is intended to mask the sounds produced by the actual taps. To further ensure this masking, pink noise is continuously played back through the headphones as background sound. GSR measures serve to quantify in real-time user's emotional arousal. The GSR shot shows the Affectiva Q Sensor (retrieved from Reuters/Affectiva/Handouts, 2012). **(C) experimental timeline.** Each experimental block lasted for 80 seconds and contained three stages: baseline1 (10 seconds) - participants only heard pink noise; feedback phase (60 seconds) - participants received real-time audio-feedback in response to their taps, while hearing pink noise; and baseline2 (10 seconds) -

participants only heard pink noise. The experimental blocks differed in the type of tapped surface (*real* or *virtual*) and in the level of strength conveyed by the tapping sounds presented as feedback (*weak*, *medium* or *strong*). At the end of each block, the subjective experience of participants during the feedback phase was assessed with a questionnaire.

Figure 2. (A) Mean perceived aggressiveness (from "tender" to "aggressive") , ability to perform the task (from "unable" to "able"), physical strength (from "weak" to "strong") and surface hardness (from "soft" to "hard"), (B) mean self-reported valence, arousal and dominance, (C) mean perceived effort (from "not at all hard to do" to "tremendously hard to do") and (D) GSR (µS) for the two surface types and three sound conditions. St = "strong", Me = "medium", We = "weak". The whiskers indicate standard error of the means. Please note that the lines connecting the points for each condition are aimed to ease visualization and do not indicate a chronological order between conditions. For GSR, change scores were calculated for each condition, by subtracting the mean response during the audio feedback period 10-65s from the mean response during the 7-8s baseline period[15].

Figure 3. (A) Mean (LOG-scores) of maximum acceleration values of tapping movements and (B) inter-tapping interval across conditions for the three phases (baseline1, baseline2 and feedback phase). (C) Acceleration changes from baseline1 across time for the real surface conditions and (D) for the virtual surface conditions. St = "strong", Me = "medium", We = "weak". The whiskers indicate standard error of the means. Please note that the lines connecting the points for each condition are aimed to ease visualization and do not indicate a chronological order between conditions.

Figure 1

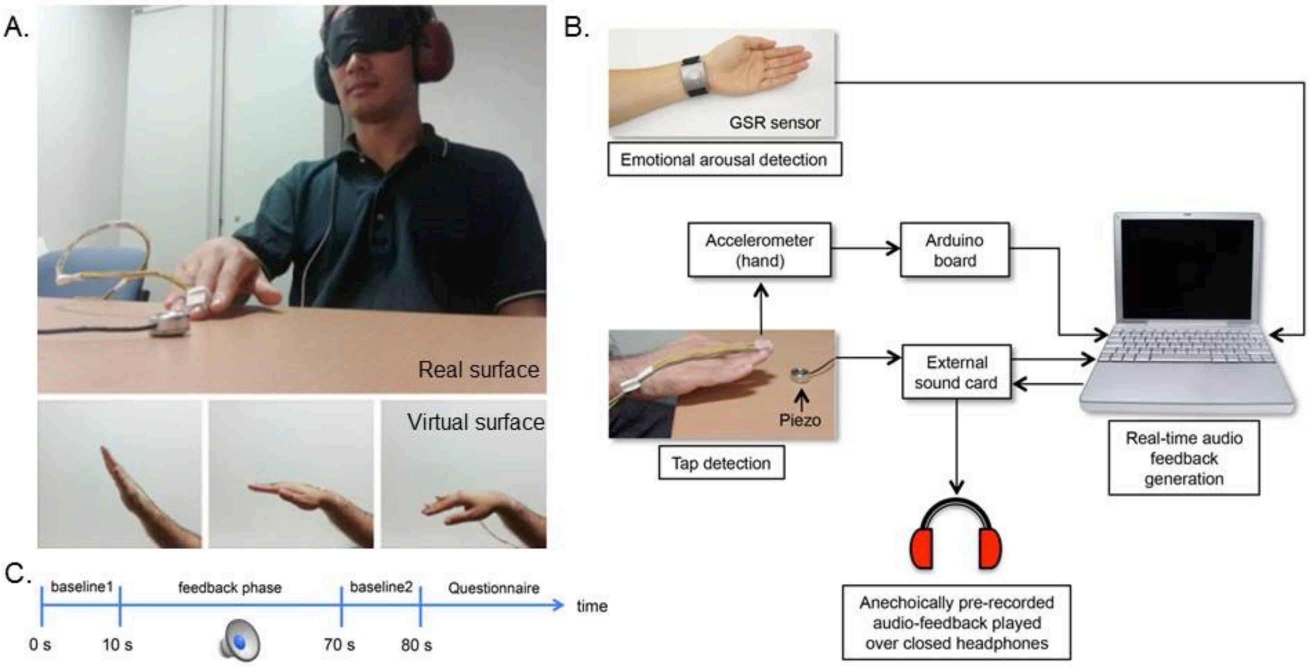

A.
Real surface
Virtual surface
B.
GSR sensor
Emotional arousal detection
Accelerometer (hand)
Arduino board
Piezo
Tap detection
External sound card
Real-time audio feedback generation
Anechoically pre-recorded audio-feedback played over closed headphones
C.
baseline1
feedback phase
baseline2
Questionnaire
time
0 s
10 s
70 s
80 s

Figure 2

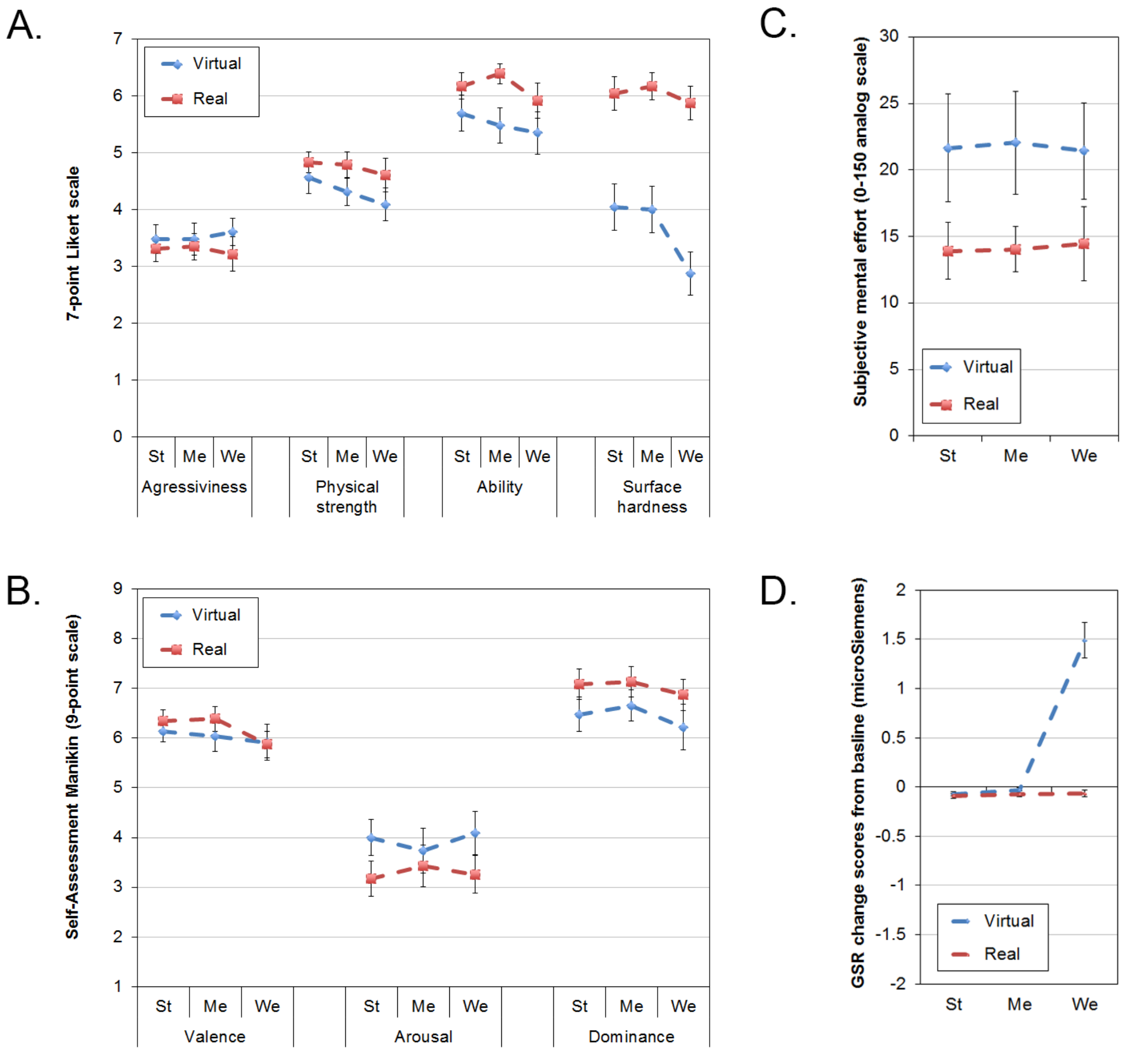
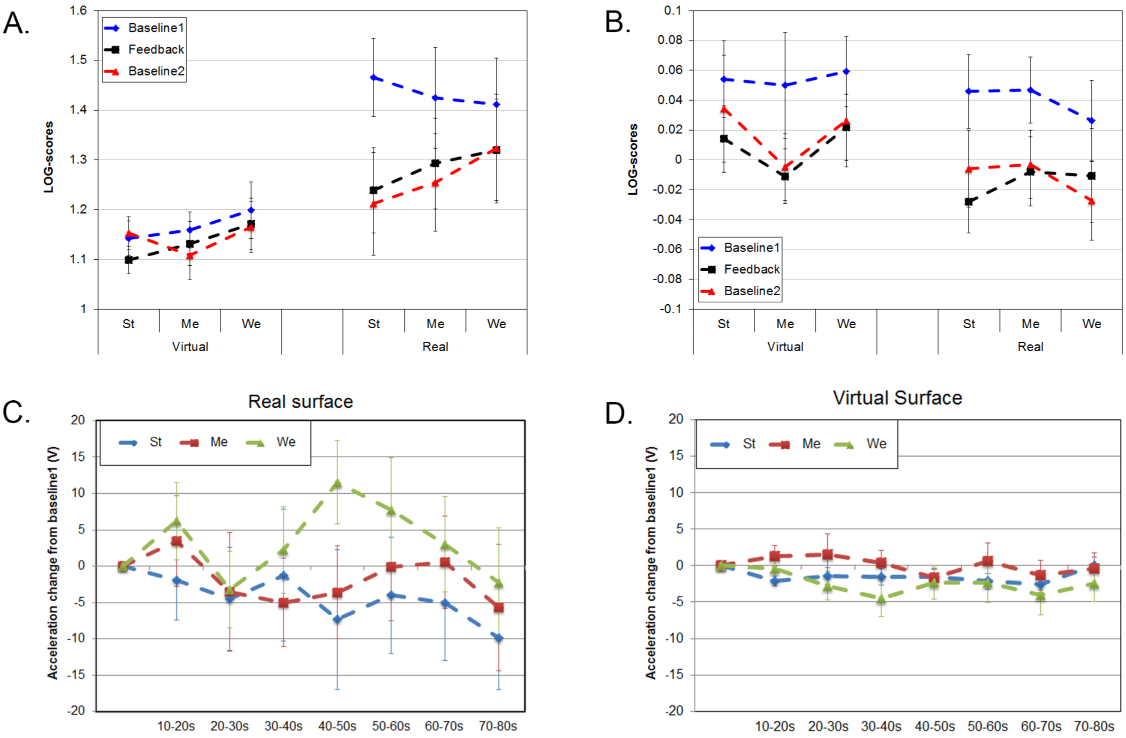

A.
7-point Likert scale
Virtual
Real
0
1
2
3
4
5
6
7
St
Me
We
Agressiviness
Physical strength
Ability
Surface hardness
B.
Self-Assessment Manikin (9-point scale)
Virtual
Real
1
2
3
4
5
6
7
8
9
Valence
Arousal
Dominance
C.
Subjective mental effort (0-150 analog scale)
0
5
10
15
20
25
30
D.
GSR change scores from basline (microSiemens)
-2
-1.5
-1
-0.5
0
0.5
1
1.5
2

Figure 3


A.
LOG-scores
1.6
1.5
1.4
1.3
1.2
1.1
1
Baseline1
Feedback
Baseline2
St
Me
We
Virtual
St
Me
We
Real
B.
LOG-scores
0.1
0.08
0.06
0.04
0.02
0
-0.02
-0.04
-0.06
-0.08
-0.1
Baseline1
Feedback
Baseline2
St
Me
We
Virtual
St
Me
We
Real
C.
Real surface
Acceleration change from baseline1 (V)
20
15
10
5
0
-5
-10
-15
-20
St
Me
We
10-20s
20-30s
30-40s
40-50s
50-60s
60-70s
70-80s
D.
Virtual Surface
Acceleration change from baseline1 (V)
20
15
10
5
0
-5
-10
-15
-20
St
Me
We
10-20s
20-30s
30-40s
40-50s
50-60s
60-70s
70-80s