

Article

Multimodal Interaction of Contextual and Non-Contextual Sound and Haptics in Virtual Simulations

Mohammed Melaisi ¹, David Rojas ² , Bill Kapralos ^{1,*}, Alvaro Uribe-Quevedo ¹  and Karen Collins ³

¹ Software Informatics Research Centre, University of Ontario Institute of Technology, Oshawa, ON L1H 7K4, Canada; mohammed.melaisi@uoit.net (M.M.); Alvaro.Quevedo@uoit.ca (A.U.-Q.)

² Faculty of Medicine, University of Toronto, Toronto, ON M5S 1A1, Canada; david.rojasgualdron@gmail.com

³ Department of Communication Arts, University of Waterloo, Waterloo, ON N2L 3G1, Canada; collinsk@uwaterloo.ca

* Correspondence: bill.kapralos@uoit.ca; Tel.: +1-905-721-8668 (ext. 2882)

Received: 24 September 2018; Accepted: 20 November 2018; Published: 26 November 2018



Abstract: Touch plays a fundamental role in our daily interactions, allowing us to interact with and perceive objects and their spatial properties. Despite its importance in the real-world, touch is often ignored in virtual environments. However, accurately simulating the sense of touch is difficult, requiring the use of high-fidelity haptic devices that are cost-prohibitive. Lower fidelity consumer-level haptic devices are becoming more widespread, yet are generally limited in perceived fidelity and the range of motion (degrees of freedom) required to realistically simulate many tasks. Studies into sound and vision suggest that the presence or absence of sound can influence task performance. Here, we explore whether the presence or absence of contextually relevant sound cues influences the performance of a simple haptic drilling task. Although the results of this study do not show any statistically significant difference in task performance with general (task-irrelevant) sound, we discuss how this is a necessary step in understanding the role of sound on haptic perception.

Keywords: multimodal interactions; haptics; sound; virtual simulation; task performance; virtual environment

1. Introduction

The simulation of the sense of touch falls under the field of haptics, which refers to machine touch and human–machine touch interactions. Haptics includes all aspects of information acquisition and object manipulation through touch by humans, machines, or a combination of the two within real, virtual, or teleoperated scenarios [1]. Haptic devices provide touch feedback that is perceived by applying tactile and/or kinesthetic stimuli to sense and/or manipulate objects with a user input device, allowing us to interact with virtual or tangible objects [2,3]. Despite the importance of the sense of touch in the real world, touch is often ignored in virtual environments (e.g., video games, virtual reality, and simulations), where historically the emphasis has been placed on reproducing—with high fidelity—the visual and (to a lesser degree) the auditory scene, while typically ignoring the other senses. Although simplistic haptic elements (notably vibration) have been used in virtual environments since at least 1996, their use is typically crude, and reviews of video games rarely mention a game’s use of haptics, indicating there is much work to be done to better employ haptics in virtual worlds.

Haptic devices can be graspable relying on force-feedback (e.g., pen-like touch devices); wearable (e.g., gloves); or touchable (e.g., touch surfaces) relying more on cutaneous feedback [4]. Accurately simulating the sense of touch is difficult as it requires providing a realistic perception of pressure,

temperature, position, movement, texture, and vibration through actuators embedded in high-end haptic devices that are cost-prohibitive for the majority of applications (i.e., upwards of \$10,000 US; see [5]). Furthermore, the cost of a haptic device is proportional to its haptic fidelity (i.e., higher fidelity implies higher price) [6]. We define *fidelity* here as the extent to which the virtual environment emulates the real-world [7]; that is, the degree of similarity between the virtual world and the real-world scenario being simulated [8]. High-end haptic devices have recently advanced and become more prominent in teleoperation to allow surgeons to perform medical procedures using robotic limbs from a distance, or teleoperators to repair components of the International Space Station [4]. Likewise, force-feedback haptic devices are now an important part of various training simulations, such as delicate medical procedures where motor skills and tactile feedback are necessary [9].

Haptic feedback has become widely available at the consumer level through actuators providing simple vibration feedback in such devices as video game controller “rumble packs”, silent vibrate settings in mobile phones, and remote-controlled robotic tools. However, the capacity of these consumer-level haptic devices to provide force feedback is limited to the small vibrational forces provided by the embedded actuators employed by the devices. As interest in haptics grows, the application of lower fidelity consumer-level haptic devices capable of providing force-feedback beyond simple vibrations is becoming more widespread, given their decreasing cost (currently approximately \$250–600 US). However, these consumer-level haptic devices are generally restrictive and cannot provide the higher level of fidelity and the range of motion required to simulate many tasks realistically [3]. Nevertheless, these lower-end devices are proving to be effective in entertainment, design-based, and educational applications, where higher levels of fidelity may not necessarily be required, and such devices have been employed in several research applications (e.g., rehabilitation, medical procedures and teleoperation) [10].

Haptics is a relatively new field, and considerable work remains to understand its use in virtual environments and training-based simulations. We do not know, for instance, how much haptic fidelity is required for learning different types of tasks. While we can assume that the cost of high-end haptic devices will drop over time, the cost of development time for simulations that incorporate these devices is likely to increase with growth in fidelity, since more accurate simulations take longer to develop and test. We do not know if such increases in cost are beneficial to the end user in terms of enhancing their task performance or learning capability.

Prior research has shown that, in multimodal virtual and real environments, other modalities can influence the user’s perception, and ultimately the performance of a task (see below). The presence or absence of sound has been shown to influence performance in, preference for, and perception of video games, for example. We are ultimately interested in how these multimodal interactions function in virtual environments, where they may be leveraged in training simulations in particular where real-life training may be undesirable, expensive or impossible (for instance, battle simulations or medical simulations). We first present an overview of the state of the current research into multimodal interactivity, and then focus specifically on haptic interactions. We then present the results of our first study into the impact that sound may have on task completion and haptic fidelity perception using a low-fidelity haptic device.

2. Background

2.1. Multimodal Interactions

In the real world, our senses are constantly exposed to stimuli from multiple sensory modalities (visual, auditory, vestibular, olfactory, and haptic) and, although the process is not exactly understood, we can integrate/process this multisensory information in our acquisition of knowledge [11]. The senses interact with one another and alter each other’s processing and perception. Decades of research has shown that vision and sound have an interactive component. For instance, the “ventriloquism effect”—the idea that we localize sound with an image, particularly if that

image is moving (see [12])—and the “McGurk effect”—visuals of conflicting syllables can alter our perception [13]—demonstrate the influence of visuals on sound. Perhaps most well-known is the Bouba–Kiki effect, in which certain visual shapes are associated with certain consonant and vowel sounds: “bouba” is generally drawn as rounded, and “kiki” as having hard angles. These associations have been demonstrated across cultures [14]. An interesting recent study has demonstrated that the Bouba–Kiki effect also occurs when the visual domain is occluded, and the participant is given shapes to touch, rather than see: in other words, sound–shape associations hold for haptic–auditory associations as well as visual [15].

Relevant to our study here is previous research into the presence or absence of sound in the performance and perception of interactive media. Jørgensen [16], as well as Lipscomb and Zehnder [17], tested the effects of having sound on and off during video game play using verbal self-reporting, and showed that sound influenced players’ perceptions of play. Robb et al. [18] found that interface sound played an important role in conveying information to game players. Psychophysiological studies have also shown sound to have an impact: Hébert et al. [19] found that playing video games with music/sound on led to higher cortisol levels than playing the same games with the sound off. Shilling et al. [20] demonstrated that playing video games with the sound on led to reductions in body temperature, but increases in heart rate and skin conductance levels compared to play with the sound off; a result also supported by Sanders and Scorgie [21]. Nacke et al., however, found neither electrodermal activity (EDA) nor facial electromyography (EMG) was influenced by the presence or absence of sounds in a game [22].

Conrad et al. [23] examined the effect of background sound (including classical music (Mozart), and metal music to induce “auditory stress”) on laparoscopic surgery. They found that metal music had a negative impact on time until task completion but did not impact task accuracy. They also found that classical music had a variable effect on time until task completion but resulted in greater task accuracy amongst all participants (laparoscopic surgeons).

Furthermore, the perception of visual fidelity can affect the perception of sound fidelity and vice versa [24]. More specifically, as described by Larsson et al. [25], if the possibilities to enhance the visuals within a virtual environment are economically or technically limited, one may consider increasing the quality of the sound channels instead. Bonneel et al. [26] observed that visual fidelity was perceived to be higher as auditory fidelity was increased, and Hulusić et al. [27] demonstrated that the addition of footstep sound effects to walking (visual) animations increased the animation smoothness perception.

Two of the authors and their colleagues previously explored the impact of background sound on various aspects of visual fidelity perception in virtual environments. One paper explored the impact of sound on visual fidelity perception with varying texture resolution [28] and found that background sound, particularly white noise, can impact our perception of high-fidelity imagery, with contextual sounds leading to an increase in fidelity perception. The authors also applied the same parameters to stereoscopic 3D and found that sound impacted visual fidelity perception again, while white noise was distracting [29]. Another study conducted with the same stimuli comparing the sound with task completion in a cel-shaded environment found that background sound impacted task performance, with white noise in particular having a negative impact on task performance [30]. On the basis of this work, then, we suspect that sound should have a similar impact on performance and perception with respect to haptic interactions.

2.2. Haptic Multimodal Interaction

In addition to the interaction of sound and visuals, there are a number of interactions between touch and other modalities that have been explored by psychologists, including the “rubber hand illusion”, which relies on an interaction between haptic and visual feedback. The rubber hand illusion has two participants: a rubber hand is placed alongside the real hand of a participant, and their real hand is covered from the participant’s view. A second participant strokes both the rubber hand and

real hand at the same time. The participant ends up seeing the rubber hand as their own, and a needle-prick in the rubber hand is read by the brain as having been applied to the real hand [31]. A similar tactile illusion known as the “parchment skin illusion” has shown that sound can bias our perception of touch (e.g., [32]). This phenomenon was reinforced by a study on audio–tactile interaction in roughness perception [33].

With respect to multimodal interaction in the virtual realm, visuals can alter the perception of stiffness of haptic feedback in virtual environments (see [34]). An emerging field of “pseudo-haptics” has evolved, which leverages haptic illusions created by multimodal interaction. For instance, visuals have been shown to influence the perception of friction, torque, size, and weight (e.g., [35–37]). Less well-studied is the interaction between auditory and haptic feedback, particularly in virtual environments. When we are used to hearing actions or touch in association with a sound, the absence of that sound can significantly impair that experience. For instance, IBM famously invented the “silent typewriter” in the 1970s but users were unsure that they had hit the keys at all, and the company added sound again (see [38]). When it comes to things that we touch with our hands in particular, sound and touch are tightly intertwined. Ballas [39] drew on studies of haptic–audio asynchrony to demonstrate that synchronous sound improves performance in a variety of tasks. Where sound is asynchronous, even to a fraction of a second, there is a marked decrease in performance (see also [40]).

McGee et al. [41] suggested that the use of sound can be employed to manipulate perception of texture in virtual devices, notably the perceived “roughness”, which has since been supported by studies into audio-haptic interactions using fMRI machines [42]. Frid et al. [43] recently described how sound in combination with haptics can influence performance in a virtual throwing task. Giordano et al. [44] established that a combination of auditory and haptic feedback on walked-upon surfaces was superior in perception to auditory or haptic feedback alone. Studies have also previously shown that the addition of audio to a haptic interface can improve performance [45].

It is our experience that auditory information is (still) often treated as an afterthought in many virtual environments and simulations (see, e.g., [46]), with music that is not contextually appropriate commonly used in casual games in particular. Likewise, sound effects that are generic samples—rather than realistic real-time simulated sounds—are still the most commonly used sounds in virtual environments, sometimes resulting in a mismatch of timings between player action and audio. For example, Collins [47] described how Nintendo Wii games fail to capture the full swing of a virtual tennis racket with a synchronized audio effect, resulting in timings that fail to synchronize. She termed these mismatches of timing “kinesonically incongruent”, and suggests that kinesonically congruent sounds may provide greater immersion and better feedback to the player. The trade-off with producing such sounds, however, is that high-fidelity recordings must be abandoned in exchange for synthesized sounds that may lack the sonic detail and fidelity of a recording. We sought to understand, therefore, whether higher auditory fidelity is preferable to higher haptic/kinesthetic fidelity, but the first step was to find a baseline as far as the influence of contextual versus non-contextual sound cues.

Despite a growing body of research demonstrating that multimodal interactions exist and have a significant influence on our perception, given the growth of haptics, there is remarkably little research into leveraging those interactions in practical applications. In particular, we are still not aware of whether the absence of one modality is always a negative influence on overall task completion and performance, usability, user experience, or perception.

3. Experimental Design

Experimental design that teases out the multitude of variables in the process of multimodal interaction is complicated. A single study cannot answer the question of whether increasing fidelity in one realm can compensate for a lower haptic fidelity device with respect to perception or task performance. Moreover, given that in most virtual environments with haptic interactions we have at least three modalities—sound, vision, and haptics—any one modality may or may not influence the other modalities. On the other hand, removal of visuals while focusing on auditory influence

on haptic performance may have different results once visuals are added back into the simulation. Therefore, to explore haptic performance, we must examine the influence of both visual and sound cues simultaneously. However, we must also compensate for redundancy to ensure that we have the correct amount of new information in any single domain while still maintaining overlap with other domains, without impeding cognitive load.

With so many factors influencing the design of multimodal interaction experiments, it is necessary that research examining multimodal interaction and its effect on task performance be undertaken in multiple stages. In our case, the first step, presented here, is to determine whether the presence or absence of contextually relevant or irrelevant sound has any impact. As the studies into games described above showed, the presence or absence of sound significantly impacted player experience, task performance, and psychophysiological response to video games. Particularly notable was that sound impacted the fidelity perception of visuals, with increased sonic fidelity resulting in an increased perception of visual fidelity.

3.1. Aim

We considered the impact of sound on haptic performance and fidelity perception by examining the effect of contextual and non-contextual sound. Our hypothesis, based on previous work, was that the presence of non-contextual sound will negatively affect task completion and haptic fidelity perception.

3.2. Methods

We developed and modeled a virtual drilling scenario that involved simulated drilling through wood with a standard household drill, in the presence of various sound cues. The scenario was developed using the Unity game engine. The drilling scenario was designed to run with two different (low-end) haptic devices: (i) the Falcon Novint Technologies, (Albuquerque, NM, USA); and (ii) the Touch 3D Stylus (3D Systems, Rock Hill, SC, USA). Such devices can simulate objects, textures, recoil, and momentum. The Falcon provides haptic feedback of 9.0 N with a resolution of 400 dpi, while the Touch 3D Stylus provides 3.3 N at 450 dpi, thus limiting the amount of realistic thrust and drilling feedback that is reproducible. It is important to note that the device's maximum exertable force occurs at the nominal (orthogonal arms) position. A high-end haptic device will typically have upwards of 40 N, indicating roughly a ten-fold increase in fidelity over the lower end devices. We collected data from drilling through an actual piece of wood and metal (with a real drill), and measuring/recording the resulting forces and vibrations (using a FLIXFORCE pressure sensor A201 connected to an Arduino UNO board). The physical drilling behaviors were simulated using a series of spring-mass systems that mimic the sensations of drilling, coded into the Unity game engine (adjusting all parameters to match the collected data) using the LibnFalcon library. The two haptic devices are illustrated in Figure 1 (Falcon) and Figure 2 (Touch 3D Stylus), while the implemented drilling models for the Falcon and Touch 3D Stylus are illustrated in Figures 3 and 4, respectively. A pilot study to test the simulation was carried out first to explore any potential training issues (described by Melaisi et al. [48]).

This particular scenario (i.e., drilling through wood) was selected because haptic feedback plays an important role during the drilling process. More specifically, it provides the user with information regarding the drilling process (e.g., how far the drill bit has traveled, the type of material being drilled, whether the drill bit becomes stuck, or whether something is wrong with the device or the drill bit). Drilling is also a fundamental component of various medical procedures, including dental surgery [49] and orthopedic surgery [50], amongst many others.



Figure 1. The Falcon haptic device.



Figure 2. The Touch 3D Stylus haptic device.

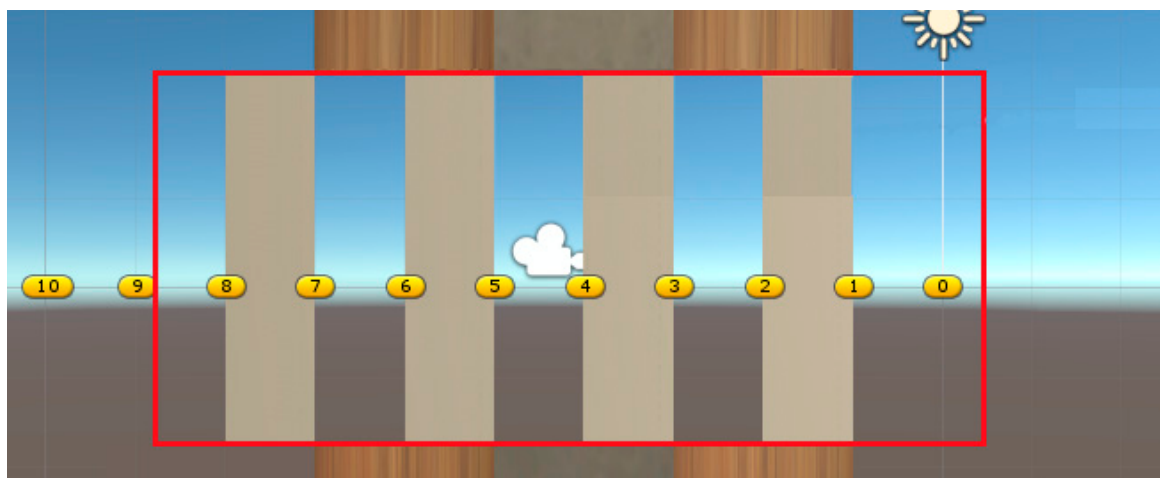


Figure 3. Spring-mass layered model for the Falcon haptic device.

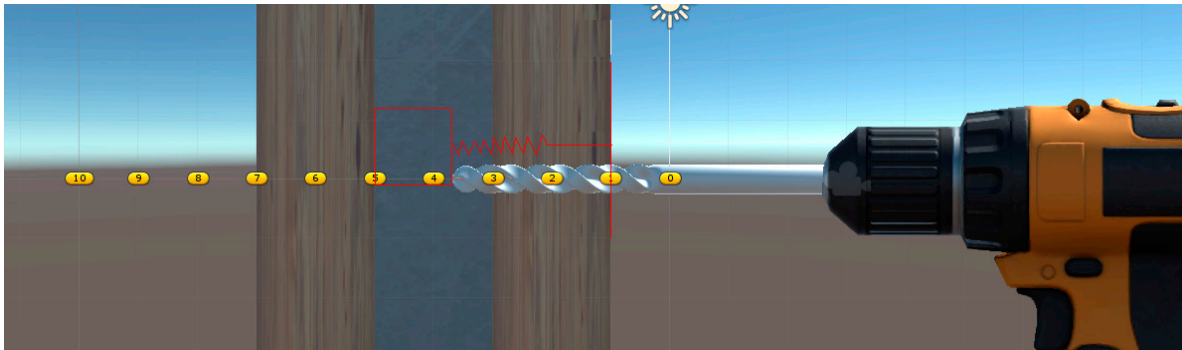


Figure 4. Spring–mass layered model for the Touch 3D Stylus haptic device.

3.3. Participants

Participants consisted of unpaid volunteer undergraduate and graduate students from the University of Ontario Institute of Technology (UOIT) in Oshawa, Canada. Six female and nine male (a total of 15) volunteers participated in the experiment. Thirteen of the participants were between 18 and 20 years old, while one was in their 30s and the other in their 40s. Before participating in the experiment, participants were asked whether they had used a drill in the past, and all answered positively. Furthermore, when asked, all of the participants reported normal hearing. Each participant was pseudo-randomly assigned to one of two experimental groups (Groups 1 and 2) using the random-stratification technique. Participants in Group 1 (eight participants) completed the training session and the intervention using the Falcon haptic device, while participants in Group 2 (seven participants) completed the training session and intervention using the Touch 3D Stylus haptic device. The experiment abided by the UOIT Research Ethics Review process for experiments involving human participants (Reference Number: 14432).

3.4. Auditory Stimuli

The selection of sounds was based on previous work described by Rojas et al. [29] where the effect of sound on visual fidelity was examined. In this case, therefore, we have no sound, contextual sound, and two irrelevant sounds. The auditory stimuli consisted of four conditions: (i) no sound; (ii) drilling sound; (iii) white noise; and (iv) classical music (“Sarabande” by Bach).

The drilling sounds corresponded to the two different materials considered in the experiment (wood and metal), and were actual recordings made of a Black and Decker household drill drilling through wood and metal (steel). The drill recordings were made in an Eckel audiometric room to limit irrelevant noise and reverberation of the generated sounds within the environment. The sounds started and stopped with the start and stop of the drill, and a looped drilling sound played for the duration of drilling. The drilling sound was contextually relevant (i.e., there should be some sounds of a drill when drilling), and was synchronized to the actions of the user (the sounds were kinesiologically congruent). However, the sound of the drill itself did not change with the depth of material, nor did the frequency of the sound change with pressure changes from the user, and thus did not provide any detailed auditory information regarding the status of their drilling beyond being on or off. In our experience, this type of sound effect would be most commonly used in simulations: fine-grained synthesis of sound in response to changes in pressure and depth would be a time-consuming task to design and would be unlikely to be used in most basic simulations and games.

All auditory stimuli were monaural and presented to the participants with a pair of headphones (AKG K240). The volume was pre-set to approximately 63 dB (roughly, normal conversation volume).

3.5. Procedure

The experiment took place within the Gaming Lab at UOIT in Oshawa, Canada (a space which houses desks, tables, and equipment including large displays), and during the time of the experiment,

only the participant and one of the experimenters were present. Participants were assigned to either Group 1 or 2 (as described above), and provided with an overview of the experiment and their required task. They were seated on a chair in front of a table that included the haptic device and a monitor that showed the simulation. Prior to the start of the experiment, participants completed a brief training session where they drilled four virtual holes aiming at a 2 cm depth, and another four at 5 cm through a piece of wood. Figure 5 provides a view of the simulation as presented to the participants in the training session. As shown, participants were provided with both a front view and a side view. In the side view, the participants viewed the drill bit going through the material (comprised on three layers: wood, metal, and wood), as they performed the drilling operation using the haptic device.

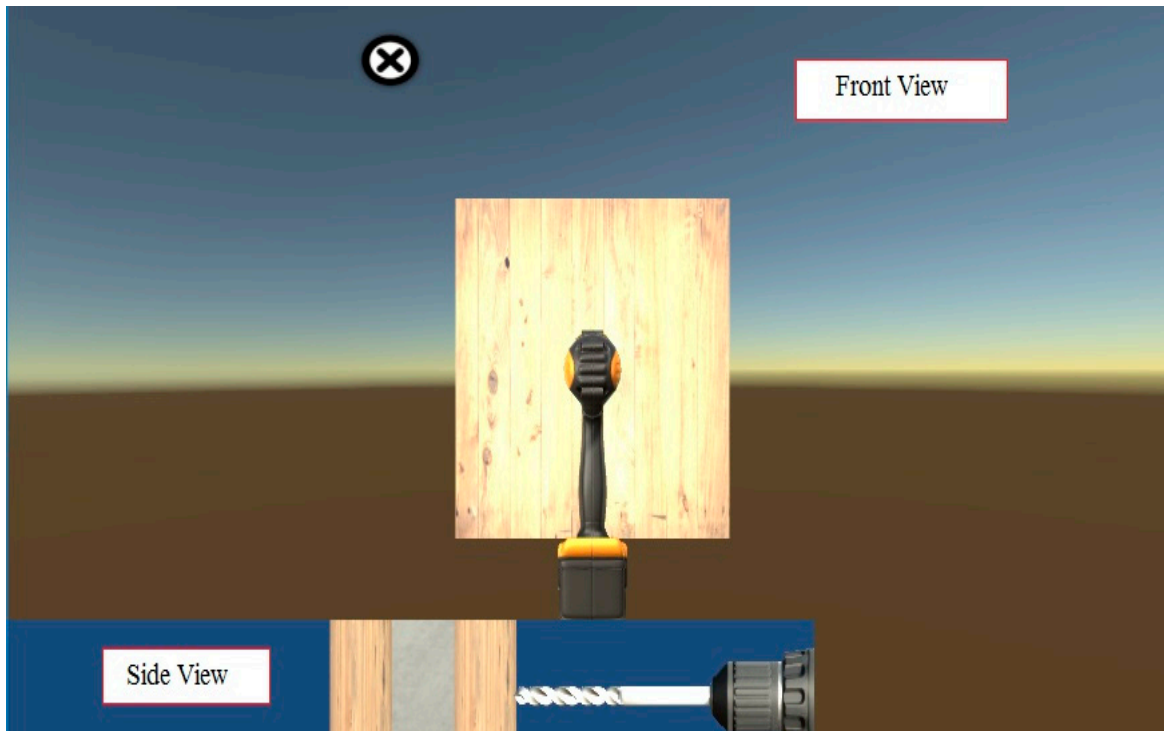


Figure 5. Participant practice session view.

After completing the training session, the participants performed drilling tasks for 24 trials. Participants were tasked with drilling a hole to a depth of either 2 cm (going through wood only), or 5 cm (going through wood and metal). During each trial, one of the four auditory conditions was presented for the duration of the trial, for eight conditions (two drilling depths \times four sound conditions), repeated three times for 24 trials. Participants had access only to the front view and thus could not see the drill depth progress through the material as provided in the side view of the practice phase. In other words, the visuals provided minimal information on the task, forcing the user to rely more and focus more on haptic feedback. To minimize possible carry-on effects, the trials were presented randomly. For each trial, we measured drilling depth accuracy and drilling task completion time.

4. Results

A MANCOVA analysis was conducted with “device” (Falcon or Touch 3D Stylus), “depth” (2 cm or 5 cm), and “auditory stimulus” (no sound, drilling sound, white noise and classical music) as independent factors (see Table 1). Two dependent variables were analyzed in this study: (i) average completion time; and (ii) drill depth accuracy. Average completion time refers to how long it took participants to complete the drilling task. Drill depth accuracy was calculated by subtracting the depth drilled from the target depth (drilled depth – target depth).

Table 1. Between-subject factors by device.

Factor	Variable	Number of Trials Run
Auditory Stimulus	No sound	60
Auditory Stimulus	Drill sound	60
Auditory Stimulus	White noise	60
Auditory Stimulus	Classical music	60
Device	Falcon	120
Device	Stylus	120
Depth	2 cm	120
Depth	5 cm	120

MANCOVA results indicate that there was a statistical significant difference for “Depth” (Roy’s Largest Root = 0.403, $F(3, 222) = 29.928$, $p < 0.01$). Univariate testing indicated that there was a significant difference in “Drill Depth Accuracy” for “Depth” ($F(1, 224) = 86.446$, $p < 0.01$). This implies that participants were more accurate at drilling when they were asked to drill to a depth of 5 cm as opposed to 2 cm, which could be accounted for by the general difficulty users had in drilling to a shallower depth overall (see Table 2). These findings are supported by the lack of significant difference of “Average Completion Time” for “Depth” ($F(1, 224) = 2.221$, $p = 0.138$), implying that participants spent about the same time when drilling to a depth of 2 cm and 5 cm. Since participants were less accurate when drilling to a depth of 2 cm (i.e., they usually drilled farther than 2 cm), one could assume that participants drilled close to 5 cm in every trial, as they spent the same amount of time drilling under the two different conditions.

Table 2. Between-subject factors by sound.

Auditory Stimulus	Depth Assigned (cm)	Mean Average Depth Drilled (cm)	Std. Deviation
No sound	2	5.74	2.33
No sound	5	5.93	2.17
Drill sound	2	5.57	2.06
Drill sound	5	6.15	2.02
White noise	2	5.63	2.18
White noise	5	6.12	2.36
Classical Music	2	5.48	2.01
Classical Music	5	5.82	1.97

There were no significant differences for “Device” (Roy’s Largest Root = 0.006, $F(3, 222) = 0.438$, $p = 0.726$), implying that the device used did not influence participants’ task performance.

The hypothesis of this study was that the presence of task-irrelevant sound should negatively affect task performance. However, there was also no significant difference for “Auditory Stimulus” (Roy’s Largest Root = 0.10, $F(3, 222) = 0.796$, $p = 0.512$).

Our results indicate that the auditory stimulus had no significant impact on task performance with respect to depth accuracy. In other words, contextually relevant sound had no significant influence on task performance in regards to drilling accuracy.

5. Discussion and Conclusions

In this paper, we have presented the results of a small, preliminary experiment that examined the influence of contextual and non-contextual sound on haptic performance through a virtual drilling task. As described, this is just the first stage in developing a better understanding of audio–haptic interaction with the aim of ultimately coupling the recently available consumer grade low-fidelity haptic devices with appropriate sound cues to provide higher haptic fidelity experience.

After analyzing the results, we found that there was no statistical significance with respect to measuring task performance with either no sound or any contextual or non-contextual sound.

These results contradicted our hypothesis that the inclusion of task-irrelevant, non-contextual sound would have a negative result on performance. However, we believe that due to the relative simplicity of the task these results indicate that the task was not complex enough for the added sound to negatively impact cognitive load. A more complex or unfamiliar task—or a task performed under more stressful conditions—will possibly yield different results.

The next stage in our process is to introduce task-relevant, kinesonically congruent sound. As stated above, the drill sound (for both wood and metal) that was used was timed (as were the other sounds) to begin and end with the drilling process, but did not provide feedback to the user regarding the drilling progress or respond sonically to pressure placed by the user. When drilling in most surfaces, a user typically has to apply more force at the start of the drilling process, because there is a barrier that must be broken before the softer wood/material is encountered (e.g., drywall, MDF, etc.), and the sound of the drill changes when that surface is broken. Likewise, the sound of a drill changes depending on its depth in the material, according to that material. This feedback is very useful to users, since one cannot easily see how deep the drill bit has gone into a material. Ideally, a change in the force/speed applied to the drill should result in a change in the sound, responding to the user. Since we aim to ultimately test the scenario in fine medical training, a drill's depth is nearly always obscured from view, and sound and haptics becomes particularly important. However, as described above, there is a trade-off between auditory fidelity in regard to the use of synthesis rather than sampled audio. It remains to be seen if the higher correspondence in terms of auditory response to user interaction would be beneficial or detrimental to the performance or perception. Therefore, task-relevant contextual sound must account for these changes in order to test whether the addition of task-relevant audio has any real impact on task performance in the tested scenario.

Given the lack of significant difference between the two lower-end devices, to test for fidelity, we must compare the results with a higher-end device, or a real drill. Finally, introducing longer-term studies may show a different influence on task learning and accuracy. Working memory—the short-term memory involved in learning a task and immediately putting it to use—is distinct from long-term memory, and task completion studies such as the one undertaken here test only working memory. Learning involves not just short-term working memory but retaining the information over longer periods. Thus, here, we only tested one type of knowledge transfer: follow-up studies after a period of several days or weeks could lead to different results.

The design and execution of studies into multimodal interactions is a complex process that involves multiple variables and multi-stage studies. A single study such as that presented here may only be able to touch the tip of the iceberg of an overall problem such as the one we propose to explore. Nevertheless, we believe that the possibilities of leveraging multimodal interactions in multimedia will, in the end, save considerable development time in designing and constructing virtual training simulations and, likely, will find use in more entertainment-based virtual environments (i.e., video games).

Author Contributions: All authors contributed substantially to the work described in this paper. Conceptualization, B.K., A.U.-Q., and K.C.; Software, M.M.; Validation, M.M., and D.R.; Formal Analysis, D.R., and M.M.; Writing—Original Draft Preparation, B.K., K.C., and A.U.-Q.; Writing—Review and Editing, K.C., B.K., and A.U.-Q., D.R.; Supervision, B.K., A.U.-Q., and K.C.; Project Administration, B.K.; Funding Acquisition, B.K., K.C., and A.U.-Q. The work described in this paper forms the basis of M.M.'s Master's thesis.

Funding: This work was financially supported by the Social Sciences and Humanities Research Council of Canada (SSHRC), Interactive and Multi-Modal Experience Research Syndicate (IMMERSe) research network, and a Natural Sciences and Engineering Research Council of Canada (NSERC) Discovery grant to B. Kapralos and A. Uribe-Quevedo.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Hayward, V.; Astley, O.R.; Cruz-Hernandez, M.; Grant, D.; Robles-De-La-Torre, G. Haptic interfaces and devices. *Sens. Rev.* **2004**, *24*, 16–29. [[CrossRef](#)]
2. Burdea, G.; Coiffet, P. *Virtual Reality Technology*; John Wiley & Sons: New York, NY, USA, 2003.
3. Jafari, N.; Adams, K.D.; Tavakoli, M. Haptics to improve task performance in people with disabilities: A review of previous studies and a guide to future research with children with disabilities. *J. Rehabil. Assist. Technol. Eng.* **2016**, *3*, 1–13. [[CrossRef](#)]
4. Culbertson, H.; Schorr, S.B.; Okamura, A.M. Haptics: The present and future of artificial touch sensation. *Annu. Rev. Control Robot. Auton. Syst.* **2018**, *1*, 385–409. [[CrossRef](#)]
5. El Saddik, A.; Orozco, M.; Eid, M.; Cha, J. Springer Series on Touch and Haptic Systems. In *Haptic Technologies: Bringing Touch to Multimedia*; Springer: New York, NY, USA, 2011.
6. Coles, T.R.; Meglan, D.; John, N.W. The role of haptics in medical training simulators: A survey of the state of the art. *IEEE Trans. Haptics* **2011**, *4*, 51–66. [[CrossRef](#)] [[PubMed](#)]
7. Andrews, D.H.; Carroll, L.A.; Bell, H. The future of selected fidelity in training devices. *Educ. Technol.* **1996**, *35*, 32–36.
8. Cook, D.A.; Hamstra, S.J.; Brydges, R.; Zendejas, B.; Szostek, J.H.; Wang, A.T.; Erwin, P.J.; Hatala, R. Comparative effectiveness of instructional design features in simulation-based education: Systematic review and meta-analysis. *Med. Teach.* **2013**, *35*, 867–898. [[CrossRef](#)] [[PubMed](#)]
9. Webster, R.; Haluck, R.S.; Zoppetti, G.; Benson, A.; Boyd, J.; Charles, N.; Reeser, J.; Sampson, S. A haptic surgical simulator for laparoscopic cholecystectomy using real-time deformable organs. In Proceedings of the IASTED International Conference Biomedical Engineering, Salzburg, Austria, 25–27 June 2003.
10. Nguyen, M.; Melaisi, M.; Cowan, B.; Uribe-Quevedo, A.; Kapralos, B. Low-end haptic devices for knee bone drilling in a serious game. *World J. Sci. Technol. Sustain. Dev.* **2017**, *14*, 241–253. [[CrossRef](#)]
11. Seitz, A.R.; van Wassenhove, V.; Shams, L. Simultaneous and independent acquisition of multisensory and unisensory associations. *Perception* **2007**, *36*, 1445–1453. [[CrossRef](#)] [[PubMed](#)]
12. Choe, C.S.; Welch, R.B.; Gilford, R.M.; Juola, J.F. The “ventriloquist effect”: Visual dominance or response bias? *Percept. Psychophys.* **1975**, *18*, 55–60. [[CrossRef](#)]
13. McGurk, H.; MacDonald, J. Hearing lips and seeing voices. *Nature* **1976**, *264*, 746–748. [[CrossRef](#)] [[PubMed](#)]
14. Köhler, W. *Gestalt Psychology*, 2nd ed.; Liveright Publishing: New York, NY, USA, 1947.
15. Fryer, L.; Freeman, J.; Pring, L. Touching words is not enough: How visual experience influences haptic–auditory associations in the “Bouba–Kiki” effect. *Cognition* **2014**, *132*, 164–173. [[CrossRef](#)] [[PubMed](#)]
16. Jørgensen, K. Left in the dark: Playing computer games with the sound turned off. In *From Pac-Man to Pop Music: Interactive Audio in Games and New Media*; Collins, K., Ed.; Ashgate Publishing Group: Aldershot, UK, 2008; pp. 163–176.
17. Lipscomb, S.D.; Zehnder, S.M. Immersion in the virtual environment: The effect of a musical score on the video gaming experience. *J. Physiol. Anthropol. Appl. Hum. Sci.* **2004**, *23*, 337–343. [[CrossRef](#)]
18. Robb, J.; Garner, T.; Collins, K.; Nacke, L. The impact of health-related sounds on player experience. *Simul. Gaming* **2017**, *48*, 402–427. [[CrossRef](#)]
19. Hébert, S.; Béland, R.; Dionne-Fournelle, O. Physiological stress response to video-game playing: The contribution of built-in music. *Life Sci.* **2005**, *76*, 2371–2380. [[CrossRef](#)] [[PubMed](#)]
20. Shilling, R.; Zyda, M.; Wardynski, E.C. Introducing emotion into military simulation and videogame design: America’s army: Operations and VIRTE. *Game-On* **2002**.
21. Sanders, R.; Scorgie, R. The Effect of Sound Delivery Methods on the User’s Sense of Presence in a Virtual Environment. Master’s Thesis, Naval Postgraduate School, Monterey, CA, USA, 2002. (Unpublished).
22. Nacke, L.E.; Grimshaw, M.N.; Lindley, C.A. More than a feeling: Measurement of sonic user experience and psychophysiology in a first-person shooter game. *Interact. Comput.* **2010**, *22*, 336–343. [[CrossRef](#)]
23. Conrad, C.; Konuk, Y.; Werner, P.; Cao, C.G.; Warshaw, A.; Rattner, D.; Jones, D.B.; Gee, D. The effect of define auditory conditions versus mental loading on the laparoscopic motor skill performance of experts. *Surg. Endosc.* **2010**, *24*, 1347–1352. [[CrossRef](#)] [[PubMed](#)]
24. Storms, S.L.; Zyda, M.J. Interactions in perceived quality of auditory-visual displays. *Presence Teleop. Virt.* **2000**, *9*, 557–580. [[CrossRef](#)]

25. Larsson, P.; Västjäll, D.; Kleiner, M. On the quality of experience: A multi-modal approach to perceptual ego-motion and sensed presence in virtual environments. In Proceedings of the First International Speech Communications Association Tutorial and Research Workshop on Auditory Quality of Systems, Herne, Germany, 23–25 April 2003.
26. Bonneel, N.; Suied, C.; Viaud-Delmon, I.; Drettakis, G. Bimodal perception of audio-visual material properties for virtual environments. *ACM Trans. Appl. Percept.* **2010**, *7*, 1–16. [[CrossRef](#)]
27. Hulusic, V.; Debattista, K.; Aggarwal, V.; Chalmers, A. Maintaining frame rate perception in interactive environments by exploiting audio-visual cross-modal interaction. *Vis. Comput.* **2010**, *27*, 57–66. [[CrossRef](#)]
28. Rojas, D.; Kapralos, B.; Hogue, A.; Collins, K.; Cristancho, S.; Nacke, L.; Conati, C.; Dubrowski, A. Developing effective serious games: The effect of ambient auditory conditions on visual fidelity perception in stereoscopic 3D. *IEEE Trans. Syst. Man Cybern. B* **2013**, *43*, 1572–1583. [[CrossRef](#)]
29. Rojas, D.; Cowan, B.; Kapralos, B.; Collins, K.; Dubrowski, A. The effect of contextual sound cues on visual fidelity perception. In Proceedings of the Medicine Meets Virtual Reality, Manhattan Beach, CA, USA, 20–22 February 2013.
30. Lam, J.; Kapralos, B.; Kanev, K.; Collins, K.; Hogue, A.; Jenkin, M. The effect of sound on visual realism perception and task completion time in a cel-shaded serious gaming virtual environment. In Proceedings of the Qomex 2015 International Workshop on Quality of MultiMedia Experience, Costa Navarino, Greece, 26–29 May 2015.
31. Ehrsson, H.H.; Spence, C.; Passingham, R.E. That’s my hand! Activity in premotor cortex reflects feeling of ownership of a limb. *Science* **2004**, *305*, 875–877. [[CrossRef](#)] [[PubMed](#)]
32. Jousmaki, V.; Hari, R. Parchment-skin illusion: Sound-biased touch. *Curr. Biol.* **1998**, *8*. [[CrossRef](#)]
33. Guest, S.; Catmur, C.; Lloyd, D.; Spence, C. Audiotactile interactions in roughness perception. *Exp. Brain Res.* **2002**, *146*, 161–171. [[CrossRef](#)] [[PubMed](#)]
34. DiFranco, D.E.; Beaugard, G.L.; Srinivasan, M.A. The effect of auditory cues on the haptic perception of stiffness in virtual environments. In Proceedings of the ASME International Mechanical Engineering Congress and Exposition, Dallas, TX, USA, 17–22 November 1997.
35. Jang, I.; Lee, D. On utilizing pseudo-haptics for cutaneous fingertip haptic device. In Proceedings of the IEEE Haptics Symposium, Houston, TX, USA, 23–26 February 2014.
36. Lecuyer, A. Simulating haptic feedback using vision: A survey of research and applications of pseudo-haptic feedback. *Presence Teleop. Virt.* **2009**, *18*, 39–53. [[CrossRef](#)]
37. Ban, Y.; Narumi, T.; Tanikawa, T.; Hirose, M. Displaying shapes with various types of surfaces using visuo-haptic interaction. In Proceedings of the ACM Symposium on Virtual Reality Software Technology, Edinburgh, UK, 11–13 November 2014.
38. Hug, D. Hear me Interact. In *Total Interaction: Theory and Practice of a New Paradigm for the Design Disciplines*; Buurman, G.M., Ed.; Birkhuser: Berlin, Germany, 2017; pp. 202–220.
39. Ballas, J. Self-produced sound: Tightly binding haptics and audio. In Proceedings of the International Workshop on Haptic and Audio Interaction Design, Seoul, Korea, 29–30 November 2007.
40. Adelstein, B.D.; Begault, D.R.; Anderson, M.R.; Wenzel, E.M. Sensitivity to haptic-audio asynchrony. In Proceedings of the International Conference on Multimodal Interfaces, Vancouver, BC, Canada, 5–7 November 2003; pp. 73–76.
41. McGee, M.R.; Gray, P.D.; Brewster, S.A. The effective combination of haptic and auditory textural information. *Lect. Notes Comput. Sci.* **2001**, *20158*, 118–126. [[CrossRef](#)]
42. Kassuba, T.; Menz, M.M.; Roder, B.; Siebner, H.R. Multisensory interactions between auditory and haptic object recognition. *Cereb. Cortex* **2013**, *23*, 1097–1107. [[CrossRef](#)] [[PubMed](#)]
43. Frid, E.; Moll, J.; Bresin, R.; Sallnas Pysander, E. Haptic Feedback combined with movement sonification using a friction sound improves task performance in a virtual throwing task. *J. Multimodal User Interfaces* **2018**, 1–12. [[CrossRef](#)]
44. Giordano, B.L.; Visell, Y.; Yao, H.Y.; Hayward, V.; Cooperstock, J.R.; McAdams, S. Identification of walked-upon materials in auditory, kinesthetic, haptic, and audio-haptic conditions. *J. Acoust. Soc. Am.* **2012**, *131*, 4002–4012. [[CrossRef](#)] [[PubMed](#)]
45. Diaz, I.; Hernantes, J.; Mansa, I.; Lozano, A.; Borro, D.; Gil, J.J.; Sanchez, E. Influence of multisensory feedback on haptic accessibility tasks. *Virtual Real.* **2006**, *10*, 31–40. [[CrossRef](#)]
46. Collins, K. *Game Sound*; MIT Press: Cambridge, MA, USA, 2008.

47. Collins, K. *Playing with Sound*; MIT Press: Cambridge, MA, USA, 2013; pp. 33–40.
48. Melaisi, M.; Nguyen, M.; Uribe, A.; Kapralos, B. The effect of sound on haptic fidelity perception. In Proceedings of the IEEE Global Engineering Education Conference, Athens, Greece, 25–28 April 2017; pp. 714–717.
49. Kinoshita, H.; Nagahata, M.; Takano, N.; Takemoto, S.; Matsunaga, S.; Abe, S.; Yoshinari, M.; Kawada, E. Development of a drilling simulator for dental implant surgery. *J. Dent. Educ.* **2016**, *80*, 83–90. [[PubMed](#)]
50. Wang, Q.; Qin, J.; Wang, W.; Shan, J.; Zhang, J.; Liu, X.; Heng, P.-A. Haptic rendering of drilling process in orthopedic surgical simulation based on the volumetric object. In Proceedings of the IEEE International Conference on Digital Signal Processing, Singapore, 21–24 July 2015; pp. 1098–1101.



© 2018 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<http://creativecommons.org/licenses/by/4.0/>).