Creating pseudo-tactile feedback in virtual reality using shared crossmodal properties of audio and tactile feedback

Isak de Villiers Bosman (iD), Koos de Beer (iD), Theo J.D. Bothma (iD)

University of Pretoria, Pretoria, South Africa

ABSTRACT

Virtual reality has the potential to enhance a variety of real-world training and entertainment applications by creating the illusion that a user of virtual reality is physically present inside the digitally created environment. However, the use of tactile feedback to convey information about this environment is often lacking in VR applications. New methods for inducing a degree of tactile feedback in users are described, which induced the illusion of a tactile experience, referred to as pseudo-tactile feedback. These methods utilised shared properties between audio and tactile feedback that can be crossmodally mapped between the two modalities in the design of a virtual reality prototype for a qualitative usability study in order to test the effectiveness and underlying causes of such feedback in the total absence of any real-world tactile feedback. Results show that participants required believable audio stimuli that they could conceive as real-world textures as well a sense of hand-ownership to suspend disbelief and construct an internally consistent mental model of the virtual environment. This allowed them to conceive believable tactile sensations that result from interaction with virtual objects inside this environment.

Keywords: virtual reality, audio, haptic, intersensory illusion, crossmodal, pseudo-haptic, pseudo-tactile

Categories: • **Human-centred computing** ∼ **Empirical studies in interaction design**

1 INTRODUCTION

Virtual reality (VR) presents novel ways of experiencing and interacting with virtual environments (VEs) and has the potential to create multi-billion dollar market within the next five years by addressing a variety of roles in a variety of areas, such as education, marketing, and training (de Regt et al., 2020). VR research and development, however, tends to focus on the visual modality at the expense of other modalities (Kapralos et al., 2017). This paper aims to

address the lack of haptic feedback in VR by focusing on the use of what are known as "crossmodal illusions", specifically relying on the crossmodal mapping between the audio and tactile modalities (Collins & Kapralos, 2019). Methods of inducing pseudo-tactile feedback using crossmodally-mappable properties of audio and tactile are described which, using qualitative data collection and analysis, were found to rely on creating what participants perceived to be believable, i.e., potentially real-world, textures as well as the perception of hand-ownership within a VE.

2 LITERATURE REVIEW

VR aims to immerse users in a simulated environment and allows them to interact with this environment with a high degree of realism (Popescu et al., 1999). Haptic feedback refers to all feedback relating to the sense of touch (Salisbury & Srinivasan, 1997) and is divided into two categories: kinaesthetic and cutaneous, with kinaesthetic feedback referring to feedback of opposing forces sensed by the muscles and joints, while cutaneous, or tactile, feedback refers to touch-sensations, such as texture or temperature (Collins & Kapralos, 2019).

The use of effective haptic feedback in VR offers several advantages, such as being able to provide VR experiences to the visually-impaired (Kapralos et al., 2017) and added level of realism (Popescu et al., 1999). However, haptic feedback is often lacking in VR due to the costliness of haptic displays (Kapralos et al., 2017).

2.1 Crossmodal illusions

Crossmodal or intersensory illusions can be defined as phenomena where stimulation in one sensory channel causes or influences perceived stimulation in another, possibly unstimulated, channel (Biocca et al., 2001; Collins & Kapralos, 2019) and have been observed across a variety of sensory modalities, including the visual, audio, and haptic (Collins & Kapralos, 2019; Whitelaw, 2008). Of particular interest for the purposes of this article are what are known as cross-modal transfers where an unstimulated sensory channel is perceived of as being stimulated due to stimuli in another channel, in other words, a user experiences sensory feedback in the total absence of such feedback being presented (Biocca et al., 2001).

This ability for feedback in one sensory modality to influence perception in another is possible due to the fact that our senses do not work in isolation, but are integrated and processed together leading to a holistic experience of objects and environments rather than an experience which is merely the sum of what each modality delivers (Kapralos et al., 2017). One factor which may contribute to the effectiveness of such illusions relates to whether the modalities themselves share properties (Hoggan & Brewster, 2006a), such as those shared between the audio and tactile modality, which are discussed in the next section.

2.2   The mapping between audio and tactile

Audio and tactile feedback both rely on pressure and vibration stimuli and, as such, inherently share certain properties such as amplitude and frequency (Eitan & Rothschild, 2011) as well as an inherent temporal property (Hoggan & Brewster, 2006a). Hoggan and Brewster, in a set of related studies, identified four properties that can be used to share audio and tactile information crossmodally (Hoggan, 2007; Hoggan & Brewster, 2006a, 2006b):

- rhythm: by altering the rate of audio beats and vibrotactile pulses

- roughness: by altering the audio timbre (such as choosing different instrument sounds) and modulating the vibrotactile amplitude

- intensity: by altering the amplitudes of audio and vibrotactile feedback

- spatial location: by altering the spatial audio's perceived location and using different vibrotactile devices.

By asking participants to align certain audio properties to tactile properties, Eitan and Rothschild identified various associations between the audio and tactile modalities (Eitan & Rothschild, 2011):

- Pitch: higher pitch was associated with sharper, rougher, and harder tactile sensations.

- Amplitude: louder sound was associated with sharper, rougher, harder, and heavier tactile sensations.

- Timbre: the timbre of a violin was more associated with blunter, rougher, and harder tactile textures, compared to a flute.

Some of these shared properties can be explained by the inherent physical similarities between sensations in the audio and tactile modalities, such as the fact that sounds with higher pitch/frequencies have shorter and thus steeper waveforms. This physical characteristic might account for the correspondence with sharper tactile sensations (Eitan & Rothschild, 2011).

By exploiting the properties shared between these modalities and relying on a user's mental model to create audio analogies of tactile sensations, it was expected that a VR application might potentially be able to induce an illusory sense of tactile feedback, which could provide VR applications with the advantages of tactile feedback.

3   RESEARCH METHODOLOGY

In order to test various possible implementations of illusion-inducing stimuli, a within-subjects usability test was used to gather data regarding participants' experiences while being presented with different combinations of stimuli in the form of a software prototype.

Qualitative data were collected across two sessions: once alongside the usability test itself in the form of a questionnaire, direct observation, and interview and once at a later date during a focus group. The study made use of qualitative data collection and analysis methods for several reasons. Firstly, the experiential nature of presence (Slater, 2003) and crossmodal illusions (Collins & Kapralos, 2019) pointed to a research method that is capable of adequately dealing with human perception and behavior (Corbin & Strauss, 2006, Ch. 1). Requiring participants to assign arbitrary numerical values to their experience was thus considered to be an inadequate way of collecting data about participants' experiences, since different numerical values could be interpreted differently by different people (Bergstrom et al., 2017). It should be noted that quantitative data were collected using the ITC-SOPI questionnaire (Lessiter et al., 2001), but since these data are not central to the current discussion they are not discussed here. A full discussion, including the purpose of the data collected with the ITC-SOPI questionnaire can be found in Bosman (2018). Open-ended interviews also allowed participants to describe their experiences in their own words, which meant that participants could thus bring up the topic of crossmodal illusions without being "primed" to do so. It was expected that fully revealing the goal of the study would prime participants to expect pseudo-tactile feedback and thus that discussion of this feedback without priming is more indicative of a genuine illusory sensation. Thus, in order to prevent bias in their answers, participants were only told that the purpose of the test was to investigate the perception of sensory phenomena in virtual reality, with no mention of intersensory illusions, tactile feedback, or binaural audio. Prior to data collection, it was explained that the questionnaire and interview focus on collecting different types of data, but both are focused on their experiences during the test.

3.1   Participants

Two different sampling methods were used for the two data collection sessions. Purposive sampling was used for the usability testing, which allowed the opportunity to select "information rich" participants who are expected to deliver in-depth data regarding their experiences. Convenience sampling was used for the focus groups, with the focus group sample drawn from the existing sample for the usability tests based on availability.

A total of 23 participants took part in the usability test and 10 of these took part in the focus groups. In terms of education level, 11 of the 23 participants had undergraduate degrees and 12 had postgraduate degrees. For the purpose of segmentation of the focus groups (Morgan et al., 1998, Bk. 2), data were collected on participants' previous computers, gaming, and VR experience using the ITC-SOPI questionnaire (Lessiter et al., 2001). In terms of computer experience, five participants rated themselves as having basic computer experience, six with intermediate experience, and 12 as having an expert level of computer experience. Three participants noted playing computer games every day, five noted "50% or more of days", five noted "often but less than 50% of days", six noted that they play computer games "occasionally", and four noted never playing computer games. Seven participants noted having previously used a VR system.

All participants gave their informed consent for both data collection sessions and were given refreshments as a reward for participating.

3.2   Materials

The software prototype was a VR application that ran on a Windows PC. An HTC Vive was used as the HMD, although the Vive controllers were not used. Instead, a Leap Motion device mounted to the front of the HMD, as shown in Figure 1, tracked participants' hands, which allowed them to interact with virtual objects without actually touching anything. The application was created with Unity and used the Google Resonance Unity plugin to deliver binaural/spatial audio.

Figure 1: The HMD-mounted Leap Motion that was used.

The prototype was divided into nine virtual rooms and participants could "teleport" through these rooms by touching virtual floating orbs. The first room was aimed at introducing participants to the virtual environment and hands and each of the rest of the rooms representing a related set of implementation strategies to induce pseudo-tactile feedback. The introductory room contained instructions and a surface with blocks which participants could interact with and move around using their virtual hands.

With the exception of the introductory room, each virtual room was structured in a similar way: two to four virtual blocks were situated in front of the participant, two teleportation mechanisms were situated to their left and right respectively, and in some rooms, instructions

were shown in front of them, as illustrated in Figure 2. Each virtual block, referred to from this point onwards as a "tactile block", represented a single audio-visual strategy for inducing pseudo-tactile feedback. Each room was a different color on the inside, to allow participants to recall specific rooms more easily.
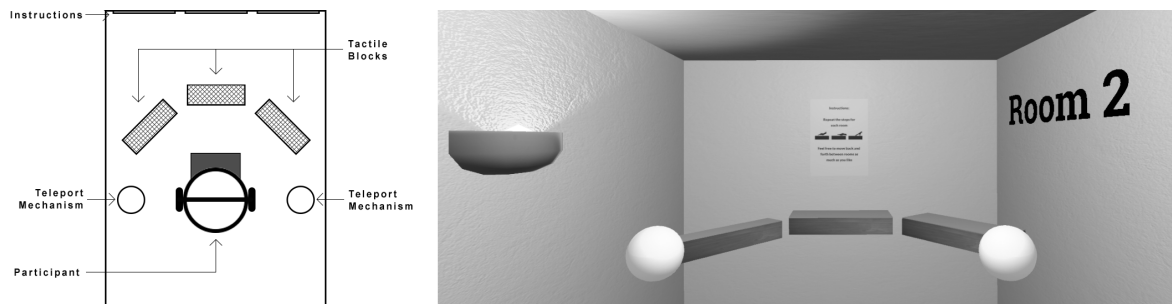

Figure 2: A top-down diagram of the layout of each virtual room (left) and a screenshot of one of the rooms (right).

Interaction with a tactile block took the form of rubbing one's hand across the surface, which produced a sound. Blocks were suspended in the air and moved down as they were pushed and moved back to their original position when released, analogous to a spring-loaded mechanism, as illustrated in Figure 3. Apart from creating a more natural interaction, this also allowed for the vertical displacement of the block to be used as a parameter for controlling various aspects of audio being played back.

With regards to audio implementations, three main strategies were adopted and combined to produce audio feedback while participants rubbed tactile blocks. The first, and simplest, was to simply play back audio recordings from rubbing textures in the real world; referred to from this point onward as the "pure texture" strategy. Two audio recordings were made, using sandpaper and a woolen jersey respectively, as these textures were expected to be recognisable and associated with contrasting textures, i.e., rough/grainy vs. soft/smooth which would make them easily distinguishable. These recordings were captured using a Line Audio CM3 microphone to record one hand rubbing across each texture in a continuous fashion, which produced a consistent rubbing sound that could be looped. Each audio file was normalised to -0.1dB for consistency between recordings and was converted to a monophonic audio file using Adobe Audition CC 2015 since Google Resonance, which localised the audio rubbing stimuli to the position of the participant's hand as they rubbed the block, works best with monophonic audio files.

All tactile blocks used a similar method for controlling the volume of audio being played back in response to being rubbed:

$$blockDisplacement = |startingBlockPositionY - currentBlockPositionY| \qquad (1)$$
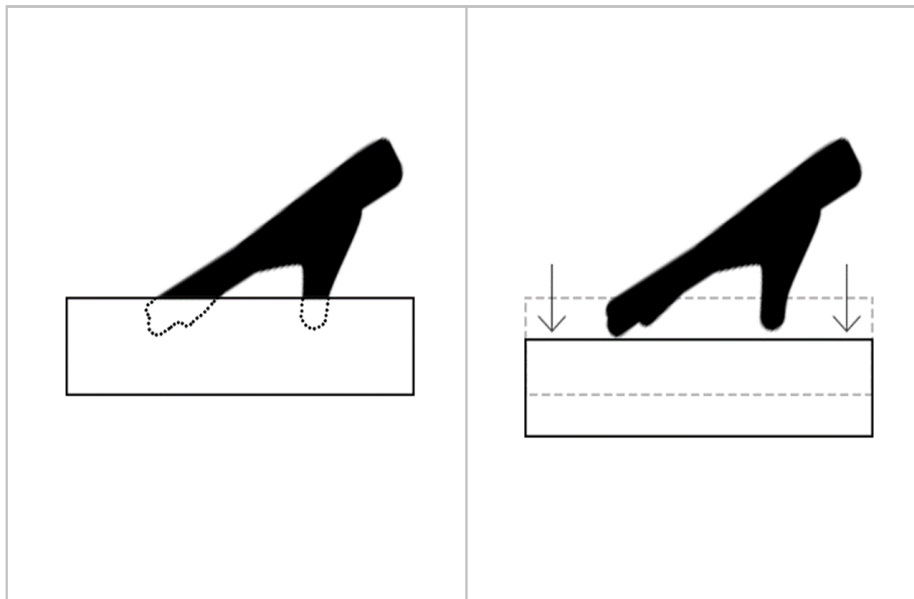
Figure 3: Interacting with a fixed block (left) would let a participant's hand go through the block, while a spring-loaded block (right) responds to the participant's hand.

$$handVelocity = \sqrt{handVelocityX^2 + handVelocityZ^2} \tag{2}$$

$$audioVolume = \frac{handVelocity}{expectedMaxHandVelocity} \times blockDisplacement \times conversionConstant \tag{3}$$

In the above formulae, the hand velocities were calculated relative to the participant's point of view, thus the x-axis referred to horisontal distance and the z-axis to depth. The use of an absolute value for calculating displacement allowed for blocks to be pushed up and down to create audio feedback. Two values above, *expectedMaxHandVelocity* and *conversionConstant*, both refer to engine-specific values determined by the researcher during informal testing to produce a natural-sounding volume. The aim for the pure texture strategy was thus to create similar behavior from the blocks as one would expect from similar interactions in the real world.

While the pure texture strategy relied on recorded, and thus "real" audio, the other two strategies relied on synthesised audio using sine-, triangle-, sawtooth-, and square-waveforms, as these are four of the most commonly used waveforms in sound synthesis (Collins, 2008, Ch. 2). Both strategies made use of a series of chromatically separated notes spanning four octaves that were created for each waveform. These two audio strategies still used the formulae above to calculate the audio volume, but both relied on crossmodal properties identified
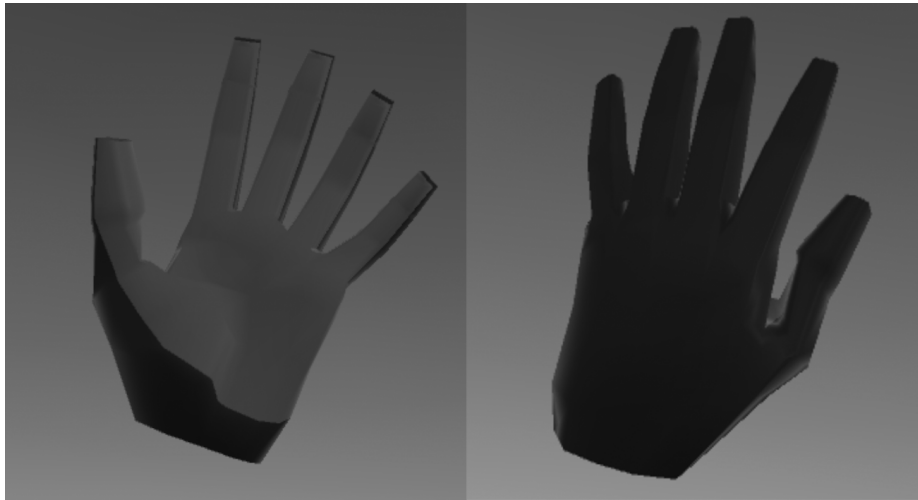
Figure 4: The front (left) and back (right) of one of the hand models used.

by Hoggan and Brewster (2006b) and Eitan and Rothschild (2011) in an attempt to create crossmodal stimuli that would map intuitively between the audio and tactile modalities. Thus, the second audio strategy, referred to from this point onward as the "atonal abstract strategy", played a single continuous note using one of the four waveforms mentioned above. As the participant rubbed their hand across the block, the note played was pitch-bent according to the following formula:

$$pitchBendAmount = \frac{handVelocity}{expectedMaxHandVelocity} \qquad (4)$$

Thus, higher hand-speeds correlated with higher-pitched sounds and the formula gives a value between zero and one due to the pitch-bending values used by Unity. For the starting note, C two octaves below middle was selected by the researcher during informal testing, as this starting note was found to be clearly audible across all waveforms. Different waveforms were selected for different blocks. In this way, the following four crossmodal properties were utilised when participants rubbed a block (Eitan & Rothschild, 2011; Hoggan & Brewster, 2006b):

- rhythm/pitch: by correlating the pitch of the audio stimulus with the movement of the participant's hand

- roughness/timbre: by varying the waveforms used across different blocks

- intensity/amplitude: by correlating the volume of the audio stimulus with the movement of the participant's hand

- spatial location: using binaural audio which localised the audio to the position of the hand as it was rubbing the block

The third audio strategy, referred to from this point onward as the "tonal abstract strategy", also used the synthesised notes across all four waveforms, but instead of playing and modifying a single note, played short notes across the different octaves. The length of the note to be played, and thus the rate of notes played, was calculated as follows:

$$audioLength = \frac{1}{handVelocity} \times lengthConstant \qquad (5)$$

Similar to the previous formulae, the lengthConstant was an engine-specific value chosen by the researcher during informal testing to resolve the final values to specific ranges. The notes to be played were selected as follows:

$$noteRange = min\left\{round\left(\frac{1}{currentBlockPositionY \times positionConstant}\right), maxOctaveRange\right\} \qquad (6)$$

$$randomNote = RandomNumber\left\{0, scaleLength\right\} \qquad (7)$$

$$randomOctave = RandomNumber\left\{0, noteRange\right\} \qquad (8)$$

$$noteToBePlayed = allNotes\left[scale[randomNote] + (randomOctaves \times 12)\right] \qquad (9)$$

The aim of the tonal abstract strategy was to create the impression of interacting with a large number of minute textural details, which provided an alternative approach to utilising the properties of rhythm/pitch and roughness/timbre (Eitan & Rothschild, 2011; Hoggan & Brewster, 2006b). As indicated in the formulae above, the selection of notes was randomised, since it was expected that a fixed musical pattern, such as an ascending scale, would be too recognisable as having musical qualities, which would make it seem unnatural. In the formulae above, *allNotes* refers to the collection of synthesised notes generated across four octaves using the four waveforms mentioned. The chromatic scale was chosen as the scale from which all blocks randomly selected notes.

Thus, the abstract strategies did not attempt to physically model audio generation from real-world interactions with objects, but rather made use of previously identified crossmodal properties shared by the audio and tactile modalities in order to create intuitively mappable audio information with the intent of inducing pseudo-tactile feedback.

With regards to the contents of the virtual rooms, these three strategies were used on their own as well as in combination with one another. A full description and motivation of the use of various strategies across all virtual rooms is listed below (with the blocks listed from left to right from the participant's point of view). Note that for the sake of brevity, the following abbreviations are used:

PT: pure texture

AA: atonal abstract

TA: tonal abstract

Table 1: The room number(s) and motivation given by each participant who noted a tactile sensation ($n = 18$) in response to the first two key questions

Room 0	Introductory room to familiarise participants with virtual hands		
Room 1	Block 1	PT: sandpaper	
	Block 2	PT: wool	
Room 2	Block 1	AA: sawtooth	
	Block 2	AA: triangle	
	Block 3	AA: sinewave	
Room 3	Block 1	TA: squarewave	
	Block 2	TA: triangle	
	Block 3	TA: sawtooth	
Room 4	Block 1	PT: wool	AA: sawtooth
	Block 2	PT: wool	TA: triangle
	Block 3	PT: wool	AA: sinewave
	Block 4	PT: wool	TA: squarewave
Room 5	Block 1	PT: sandpaper	AA: triangle
	Block 2	PT: sandpaper	TA: sawtooth
	Block 3	PT: sandpaper	AA: squarewave
	Block 4	PT: sandpaper	TA: sinewave
Room 6	Block 1	TA: sawtooth	AA: sinewave
	Block 2	TA: squarewave	AA: sawtooth
	Block 3	TA: triangle	AA: squarewave
	Block 4	TA: sinewave	AA: triangle

Room 7	Block 1	PT: wool	TA: sinewave	AA: triangle
	Block 2	PT: wool	TA: triangle	AA: sawtooth
	Block 3	PT: wool	TA: squarewave	AA: sinewave
	Block 4	PT: wool	TA: sawtooth	AA: squarewave
Room 8	Block 1	PT: sandpaper	TA: triangle	AA: sinewave
	Block 2	PT: sandpaper	TA: sinewave	AA: squarewave
	Block 3	PT: sandpaper	TA: sawtooth	AA: triangle
	Block 4	PT: sandpaper	TA: squarewave	AA: sawtooth

The intent of the experiment was not to include all possible combinations of strategies and waveforms, but rather to explore all possible approaches with regards to the strategies used and the creation of similar or contrasting textures in terms of smoothness or roughness. Thus, the goal of the software prototype was to utilise crossmodally-shared attributes of audio and tactile

feedback to implement various possible strategies for inducing pseudo-tactile feedback, which would allow for the collection of qualitative data that offer explanations as to the underlying mechanisms regarding these illusory sensations.

The procedure for data collection went as follows: each participant was given written consent forms to sign, which gave permission for the recording of various data throughout the test. The goal of the test was explained to participants as focusing on "investigating the perception of sensory phenomena in virtual reality". Participants were then aided with the equipment and given general instructions for the purpose of the study.

Once participants put on the headset, the software prototype was started, thus beginning the test. Participants were not required or expected to spend a specific amount of time inside each room; they were only instructed, via text instructions in the room, to "Feel free to move back and forth between rooms as much as you like". Room 8 instructed them to remove the headset and inform the researcher when they were satisfied that they had explored enough. During all data collection the word "experiment" was used instead of "test" to frame participation as an exploration rather than a performance in order to reduce possible anxiety.

Once the test had started, video recording started in the form of a webcam capturing participants' behavior in the real world and screen capturing software capturing their point of view in VR. After completing the questionnaire, a semi-structured interview was conducted with each participant, making use of open-ended questions. The interview script can be found in Appendix A. At a later date, two focus groups were held with different groups of participants. The focus groups were moderately structured and made use of pre-defined questions to spur conversation between participants. The focus group script can be found in Appendix B. Collected data were analysed using constant comparative analysis by coding discrete units of data using Atlas.ti.

4   RESULTS

The interviews included three key questions aimed at assessing the effectiveness of the prototype in achieving its goal of utilising crossmodally-shared attributes of audio and tactile feedback to implement various possible strategies for inducing pseudo-tactile feedback. However, to avoid priming participants early in the interview, these questions were only asked near the end of the interview with opportunity for participants to express their sensory experiences throughout. These three key questions were as follows:

1. In which room would you say what you experienced came closest to actual touch?

2. Can you elaborate on why that room's touch experience felt more real than the others?

3. How would you summarise your overall experience in terms of your sense of touch during the experiment?

Thus, it was assumed that descriptions of pseudo-tactile experiences that were brought up unprompted, i.e., before the three key questions, were more indicative of genuine illusory

experiences. From the 23 participants who took part in the study, the results as illustrated in Figure 5, in terms of descriptions of pseudo-tactile feedback, were as follows:

- 16 participants noted a tactile experience before the three key questions were asked

- 2 participants noted a tactile experience in response to the key questions

- 3 participants gave responses to questions that were considered contradicting

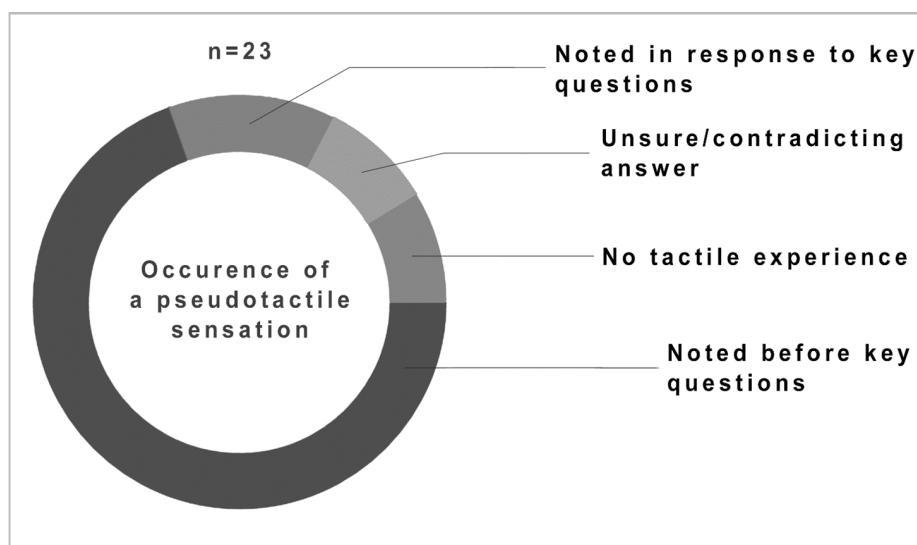- 2 participants clearly noted experiencing no tactile sensations

Figure 5: The distribution of participants' descriptions of experiencing pseudo-tactile feedback.

Interview responses were considered to indicate tactile sensations if participants expressed being able to "feel" the tactile blocks, for example:

> the sound that emerges from you touching the blocks, it creates an actual sensation in your hands, you can actually feel as if you are touching the actual blocks due to the sound.

Responses were categorised as contradictory if participants' responses both suggested being able and unable to feel something when rubbing the blocks, which were not considered as successful in the study's goal.

The pseudo-tactile sensations were also found to be robust, with 11 out of 18 participants who described a sense of touch also discussing the impossibility of this experience (without being prompted to do so), for example:

> In your brain you know it isn't there, but kind of if I'm there it feels… understand? It really feels real.

Descriptions of touch-sensations were most often described as either a touching- or tingling-sensation, the former being described in terms of touching a surface, such as

> … it felt as if I'm really touching something

and the latter in terms of a "tingling" sensation, such as

> It felt as if there was almost a sort of pins and needles in my fingers if I felt across the blocks.

Participants who experienced illusory touch discussed a variety of factors as the underlying cause. Table 1 and Figure 6 below show a breakdown of the responses to the first two key questions with responses divided into main categories listed by participants as the main influence on their illusory sense of touch: audio feedback (A), visual feedback (V), both (B), and other (O). It should be noted that the table only includes responses from participants who also indicated a successful pseudo-tactile sensation and that participants could include more than one room in their response.

Table 2: The tactile blocks in each virtual room

Participant number ($n = 18$)

Room	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18
1	V				B	A	B					B			V	A	V	
2				A														A
3																		
4			A									A						
5			A	A												A		
6		V														A		
7								V	B	V				B		A		
8					A			O	V				A					A

The remainder of this section describes themes that emerged from qualitative data analysis of interview and focus group data with regards to the causes of pseudo-tactile sensations. Two overarching themes that arose across the data were the effects of presence as well as realism and believability.

With regards to presence, participants' descriptions of feeling physically "present" inside the VE tended to align with Heeter and Allbritton's (2015) four dimensions of presence, except for social presence, since the VE did not contain any other social beings whom participants could interact with. Of particular importance were descriptions of personal presence, which most often indicated hand-ownership, i.e., the sense that the virtual representations of hands in the VE were their own hands. This manifested both explicitly, when participants noted
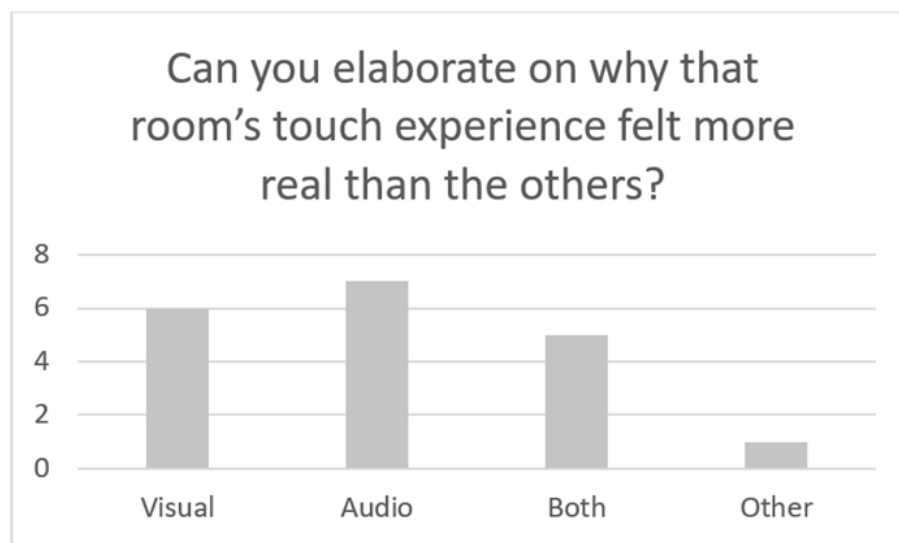
Figure 6: Broad overview of which sense participants referred to in response to the second key question. Only one participant did not specifically refer to either visual or audio feedback in their response.

experiencing that the virtual hands were their own hands ("The hands were a good touch, it made me feel like those were my real hands that can touch things."), and implicitly, when participants spoke about the virtual hand representations as if they were their own hands, ("… I could see my hands and they moved").

With regards to realism and believability, as mentioned, the interview included a key question aimed at why certain tactile blocks elicited more effective pseudo-tactile sensations than others. Referring again to the second key question ("Can you elaborate on why that room's touch experience felt more real than the others?"), 14 out of 18 participants who noted an illusory response cited either a sense of realism or believability as the key component, for example:

> … the sound when you touch [the block], I must say, it felt more as if it really could be like that. So, you know, the feeling you would get if you rubbed over a block with your hand…

However, the distinction between realism and believability was not clear, as participants often described certain unrealistic/impossible aspects of the VE, such as the movement of the blocks, as "realistic". Thus, believability, for the current discussion, refers to appearances or behaviors that participants were able to conceptualise as being possibly real in certain contexts and allows some deviation from objective reality, as opposed to true verisimilitude which requires that appearances or behaviors have a direct counterpart in objective reality. Qualitative data for the current study indicate three main themes that related to how participants experienced the VE and its contents as believable enough to experience pseudo-tactile feedback as a result:

expectation, internal consistency, and suspension of disbelief.

The first factor that contributed to the believability of the VE arose from the fact that when participants entered VR, they carried over expectations from the real world with regards to how the VE should appear and behave. However, the expectation which participants had with regards to the behavior of the blocks was not limited to that with real-world counterparts, for example, the behavior of the blocks was often described as "realistic", even though their behavior of hovering in a point in space and being "spring-bound" to that point was blatantly unrealistic. An example of this unrealistic behavior being perceived as believable due to participants' expectations is illustrated in the following exchange:

> **[Researcher]**: Can anyone think of why [they experienced pseudo-haptic feedback]?
>
> **[Participant]**: Maybe it's because you're imagining that 'ok, there's a block and I want to push it down' and by doing that it feels like 'ok, obviously if you push something down it wants to come up naturally'.

Certain aspects of believability were found to be a requirement for experiencing pseudo-tactile feedback, specifically that of believable audio textures for the blocks and believable hands. Some abstract audio strategies were perceived as "unnatural", "digital", etc., which highlighted the artificiality of interacting with the block. However, for audio strategies that could induce pseudo-tactile feedback, all participants except two who noted these sensations attributed them to the fact that the audio and/or visual feedback made their experiences of those blocks more real or easier to attribute to real-world (tactile) textures. Of the two participants that did not give this explanation, one attributed their tactile sensations to forgetting about the virtual nature of the blocks and one noted that the blocks that caused the closest touch-sensation caused the most "feedback" and made them shudder and pull back their hands. (It should also be noted that the last participant, when prompted, stated that the blocks that made them shudder and pull back their hands reminded them of sandpaper and charcoal respectively.)

Thus, being able to attribute the sense of a texture created by the audio and/or visual feedback to a real-world texture was found to be one of the primary contributors to an illusory touch experience. Where participants could not describe the audio and/or visual feedback from the blocks in terms of textures they could interact with in the real world, the blocks creating these textures were found to be ineffective in creating tactile sensations. A possible reason for this is the bilateral nature of touch (Salisbury & Srinivasan, 1997), which requires direct interaction with the physical environment to experience a sense of touch. For this reason, while audio and visual feedback both allow for the creation of "abstract" representations that do not appear in nature, such as synthesised audio, these results suggest that tactile feedback does not allow for this.

The same argument applies to the experience of believable virtual hands and hand ownership: since touch-sensations require direct contact with the physical environment, participants needed virtual hands that were believable as hands and that they experienced as "belonging"

to them, i.e., hand-ownership, in order to experience tactile sensations. The data support this argument, as all participants except one who noted pseudo-tactile sensations also explicitly or implicitly noted a sense of hand-ownership (the one exception did not refer to the virtual hands during any data collection).

The second factor that contributed to the believability of the VE was the internal consistency of the VE in terms of its behavior and attributes. Data from participants confirm findings by (Biocca et al., 2001) stating that participants require consistent, shared information between senses to construct an internally consistent mental model of a VE. Internal consistency was also found to have a reciprocal relationship with expectation, as both were found to inform the other to an extent. Expectation was found to inform consistency insofar as that participants carried over expectations from real-world interactions and that if those expectations were met, they could construct an internally consistent mental model of behaviors in the VE. Conversely, participants also relied on this mental model to inform their expectation of future behaviors in the VE, thus creating expectation of how such future behaviors should function.
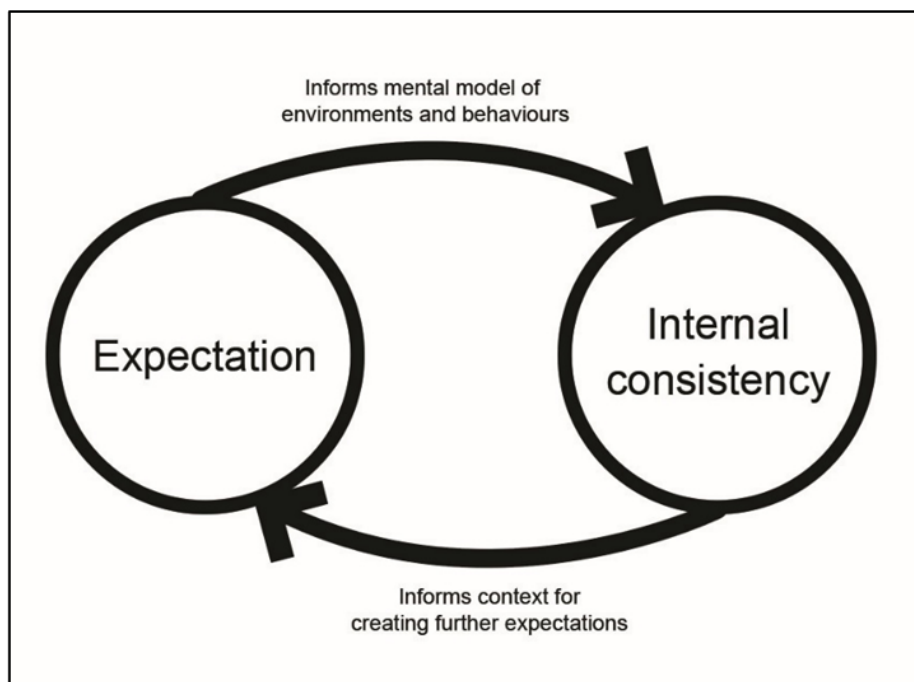

Figure 7: Expectation and internal consistency were found to have a reciprocal relationship.

Finally, a willing suspension of disbelief was found to be a necessary component in experiencing the VE as believable enough to induce pseudo-tactile feedback. Participants often described their suspension of disbelief in terms of a threshold rather than an absolute

> [I only experienced tingling sensation when] any form of, with the fingers, with the hand... [*sic*] with the palm or back of my hand, for specific things, for when it made

> sense, like with the carpet and then the sound of the carpet, I think everything
> added together and then maybe potentially warped my brain into thinking 'this
> could potentially be something' and that initiating a reaction of some sort.

In other words, participants could only suspend disbelief enough to experience illusory sensations when it was "easy enough" for them to do so given the stimuli associated with the different blocks, such as the aforementioned participant only experiencing these sensations when it "made sense".

Given the aforementioned explanation regarding the requirement for the audio and visual feedback to be associable with real-world textures, this would suggest that there was a variability in the effort with which participants were able to suspend disbelief, which was described by two participants as a "jump" or "link" that had to be made in their minds between the stimuli presented by the VE and the expected real-world equivalent. Participants were only able to experience pseudo-tactile sensations if this jump/link was sufficiently low, which required that their expectations were met and that the internal consistency of the VE was maintained.

5 CONCLUSION

This paper describes a software implementation that utilised various properties shared between the audio and tactile modalities and was able to induce pseudo-tactile sensations with moderate success. Qualitative data were collected from participants who used this software prototype, which indicated several factors which were integral in the induction of these sensations. The results of this study suggests that illusory tactile sensations can be used as a useful alternative in the absence of sensory displays that are able to produce real tactile feedback.

For participants to be able to "feel" sensations in their hands from interacting with virtual blocks, they needed to suspend disbelief in the unrealism of the VE and accept the blocks, to some degree, as believable physical objects capable of creating physical sensations. This required threshold of suspension of disbelief depended on the stimuli presented by the blocks as well as the general behavior of the blocks in response to participants' actions. If the virtual blocks behaved enough like real blocks and if the audio that resulted from rubbing them was relatable to real-world textures, participants could conceive of the tactile sensations that should arise as a result of interacting with the blocks. This also suggests that the tactile modality does not allow for the creation of "abstract" representations, such as the audio modality being able to represent synthesised sounds that do not appear in physical reality.

The same applied to the perception of hand-ownership, since participants needed enough of an experience that the virtual hands "belonged" to them in order to experience an illusory touch-sensation. Conversely, if this threshold was too high, participants were unable to suspend their disbelief enough to overlook the artificiality of the stimuli and perceive tactile sensations as a result. Thus, the three factors identified present a guide on how such feedback can be designed and where the limits of this method of inducing tactile feedback lie. By meeting users' expectations that stem from the real world and sustaining an internally consistent

set of audio-visual stimuli that allows users to suspend their disbelief in virtual objects, VR applications can create robust illusory tactile stimuli.

Further research might include investigating the individualised factors that might affect the experience of illusory tactile feedback, including demographic, psychological, and psysiological factors. Used alongside tactile displays, the audio strategies described in this study could also alter or enhance perceived tactile sensation, which presents another potential avenue for study.

References

Bergstrom, I., Azevedo, S., Papiotis, P., Saldanha, N. & Slater, M. (2017). The plausibility of a string quartet performance in virtual reality. *IEEE Transactions on Visualization and Computer Graphics, 23*(4), 1352–1359.

Biocca, F., Kim, J. & Choi, Y. (2001). Visual touch in virtual environments: An exploratory study of presence, multimodal interfaces, and cross-modal sensory illusions. *Presence: Teleoperators and Virtual Environments, 10*(3), 247–265. https://doi.org/10.1162/105474601300343595

Bosman, I. (2018). *Using binaural audio for inducing intersensory illusions to create illusory tactile feedback in virtual reality* (Master's thesis). University of Pretoria. Pretoria.

Collins, K. (2008). *Game sound: An introduction to the history, theory, and practice of video game music and sound design.* MIT Press.

Collins, K. & Kapralos, B. (2019). Pseudo-haptics: Leveraging cross-modal perception in virtual environments. *The Senses and Society, 14*(3), 313–329. https://doi.org/10.1080/17458927.2019.1619318

Corbin, J. & Strauss, A. (2006). *Basics of qualitative research.* Sage.

de Regt, A., Barnes, S. J. & Plangger, K. (2020). The virtual reality value chain. *Business Horizons, 63*(6), 737–748. https://doi.org/10.1016/j.bushor.2020.08.002

Eitan, Z. & Rothschild, I. (2011). How music touches: Musical parameters and listeners' audio-tactile metaphorical mappings. *Psychology of Music, 39*(4), 449–467. https://doi.org/10.1177/0305735610377592

Heeter, C. & Allbritton, M. (2015). Playing with presence: How meditation can increase the experience of embodied presence in a virtual world [Last accessed 21 Jun 2021]. http://www.fdg2015.org/papers/fdg2015_extended_abstract_07.pdf

Hoggan, E. E. (2007). Crossmodal interaction: Using audio or tactile displays in mobile devices. In C. Baranauskas, P. Palanque, J. Abascal & S. D. J. Barbosa (Eds.), *Human-computer interaction – interact 2007* (pp. 577–579). Springer.

Hoggan, E. E. & Brewster, S. A. (2006a). Crossmodal icons for information display. *CHI '06 Extended Abstracts on Human Factors in Computing Systems,* 857–862. https://doi.org/10.1145/1125451.1125619

Hoggan, E. E. & Brewster, S. A. (2006b). Crossmodal spatial location: Initial experiments. *Proceedings of the 4th Nordic Conference on Human-computer Interaction: Changing Roles*, 469–472. https://doi.org/10.1145/1182475.1182539

Kapralos, B., Collins, K. & Uribe-Quevedo, A. (2017). The senses and virtual environments. *The Senses and Society*, *12*(1), 69–75. https://doi.org/10.1080/17458927.2017.1268822

Lessiter, J., Freeman, J., Keogh, E. & Davidoff, J. (2001). A cross-media presence questionnaire: The ITC-sense of presence inventory. *Presence, 10*(3), 282–297. https://doi.org/10.1162/105474601300343612

Morgan, D. L., Krueger, R. A. & King, J. A. (1998). *Focus group kit* [6 vols.]. Sage.

Popescu, V., Burdea, G. & Bouzit, M. (1999). Virtual reality simulation modeling for a haptic glove. *Computer Animation, 1999. Proceedings*, 195–200. https://doi.org/10.1109/CA.1999.781212

Salisbury, J. K. & Srinivasan, M. A. (1997). Phantom-based haptic interaction with virtual objects. *IEEE Computer Graphics and Applications, 17*(5), 6–10. https://doi.org/10.1109/MCG.1997.1626171

Slater, M. (2003). A note on presence terminology [Last accessed 20 Jul 2021]. http://www0.cs.ucl.ac.uk/research/vr/Projects/Presencia/ConsortiumPublications/ucl_cs_papers/presence-terminology.htm

Whitelaw, M. (2008). Synesthesia and cross-modality in contemporary audiovisuals. *The Senses and Society, 3*(3), 259–276. https://doi.org/10.2752/174589308X331314.

A APPENDIX: INTERVIEW QUESTIONS

The interview was semi-structured, with the questions listed below. Every interview was recorded using a voice-recorder. Participants were given refreshments before the start of the interview.

The interview started with a primer, explaining the details of the interview:

> Just to reiterate, there are no right or wrong answers during this interview. The only right answers are your honest opinions and experiences during and after using the VR system.

Question 1. You indicated on the questionnaire that you have [*x*] amount of experience playing video games.[1] Can you elaborate on your answer?

Question 2. You also indicated that you have [*y*] amount of knowledge on virtual reality and virtual reality technologies.[2] Can you elaborate on that answer?

Question 3. How would you describe your experience inside the virtual environment during the experiment, especially while rubbing the blocks?[3]

---

[1] *'questionnaire'* and *'x'* referred to their answer on the ITC-SOPI questionnaire.

[2] *'y'* referred to their answer on the ITC-SOPI questionnaire.

[3] The interviewer would point out that the participant could go back to refresh their memory.

Question 4. How would you describe your experience in terms of any notable sensory experiences in the first room?

Question 5. How would you describe your experience in terms of any notable sensory experiences in the [*second*] room?[4]

Question 6. In which room would you say what you experienced came closest to actual touch?

Question 7. Can you elaborate on why that room's touch experience felt more real than the others?

Question 8. How would you summarize your overall experience in terms of your sense of touch during the experiment?

Question 9. Did you experience anything else that we haven't discussed that you think is noteworthy and would like to add?

The interview concluded with the prompt

> If you remember anything else, please let me know,

after which the researcher's contact information was provided to the participant.

B APPENDIX: FOCUS GROUP QUESTIONS

The interview made use of a pre-established structure consisting of a set of questions, as listed below. Most of the session was voice-recorded starting from question 2, so as to maintain confidentiality. Participants were given refreshments before and after the focus group session as their reward.

Both focus groups started with a primer, explaining the purpose and details of the focus group:

> Same as the interviews, there are no right or wrong answers during this discussion.
> The purpose is to get everyone's opinions and experiences in their own words.

For the introductory question, the two groups were asked slightly different questions. The aim of this was to establish rapport between participants that play games by drawing attention to their collective gaming-experience.

Question 1, Group 1. Can everyone please introduce themselves by giving their names, what they are studying or have studied in the past, and their favorite video game at the moment?

Question 1, Group 2. Can everyone please introduce themselves by giving their names and what they are studying or have studied in the past?

[*Voice recording started at this point.*]

Question 2. Explain your feelings toward virtual reality. Are you interested, disinterested, excited, apathetic, etc.?

---

[4]This question was repeated for each virtual room.

Question 3. It is widely accepted that a sense of touch is important for making sense of the real world, but how do you think virtual reality would benefit from being able to provide a sense of touch, i.e., a tactile experience?

Question 4. For the next couple of questions, I want to discuss a few topics that kept coming up in the interviews, the first being realism. [*A number of sub-topics of realism were discussed, including the importance of realism, the effect of unrealistic interactions, the effect of acting on the environment on perceived realism, and the malleability of realism.*]

Question 5. The next topic is that of the role of the environment: did anyone experience different effects or feelings in the different rooms because of the room colors? Such as temperature?

Question 6. Another topic that came up during the interviews was a sense of pressure when pushing the blocks. Did anyone experience something like this?

Question 7. Some of you also noted feeling a sort of a sense of touch during the experiment, can anyone elaborate on this? Can anyone elaborate on the nature of this "sense-of-touch" experience?

Question 8. Those who did feel a sense of touch, can you elaborate on where it was and what made that room or those block different?

Question 9. During the interview, when discussing blocks and their differences, most of you used textural descriptive words to describe blocks, even those who didn't "feel" anything. Can anyone elaborate on the reason for this?

Question 10. Was there anything specific in the virtual environment that diminished your experience or kept you from feeling that you were part of the virtual world?[5]

Question 11. [*During the discussions, the researcher kept track of the main points and briefly summarize them. At this point, this summary was read out to all participants.*] Does [*the summary*] adequately summarize everyone's experiences?

Question 12. Have I overlooked anything noteworthy or interesting that you still want to discuss?

[*Participants were thanked for their time and given the remainder of the refreshments.*]

---

[5]The following were brought up: "digital" sounds, inconsistent behavior of hands, specific rooms, and colors and lighting.