Can you hear the Colour? Towards a Synaesthetic and Multimodal Design Approach in Virtual Worlds

Victoria Wright
School of Informatics
University of Leicester, University Road
Leicester, LE1 7RH, UK
vrv3@leicester.ac.uk

Genovefa Kefalidou
School of Informatics
University of Leicester, University Road
Leicester, LE1 7RH, UK
gk169@leicester.ac.uk

Synaesthesia is a phenomenon where senses naturally combine resulting in, for example, ‘seeing’ music or ‘hearing’ colours. It is of interest in the field of Human-Computer Interaction as a way of creating new or enhanced experiences and interactions with Mixed Reality technologies. In Virtual Reality, research has mainly focused on evaluating advanced graphics and capturing immersion levels and User Experience within ‘typical’ and ‘expected’ interactions. This paper investigates how multimodal design characteristics can lay the foundations to a more ‘synaesthetic’ design approach in Mixed Reality to identify how ‘atypical’ interactions can also affect User Experience. 20 participants completed a maze activity, emotion and immersion surveys and interviews. Results suggest a significant increase in surprise, pride and inspiration and a decrease in interest and enthusiasm. The visual and audio aspects were well received by participants and the sensory elements had a positive effect on User Experience. Time perception was measured and 90 per cent of participants’ time estimations were longer than the actual time. Change blindness was investigated with most participants not noticing the visual or audio changes. Finally, we discuss how this study can inform future projects which aim to implement a synaesthetic-oriented and multimodal approach in Mixed Reality design.

Multisensory. Virtual Environment. Synaesthetic-oriented approach. Change blindness. Time perception.

1. INTRODUCTION

In recent years, Mixed Reality (MR) technologies have become more advanced and more prominent in fields of healthcare (McLay et al., 2014; Striem-Amit, Guendelman & Amedi, 2012), commerce (Van Kerrebroeck, Brengman & Willems, 2017), as well as leisure, with Virtual Reality (VR) headsets such as the Oculus Rift (Oculus, 2019) and Augmented Reality (AR) games such as Pokémon Go (Niantic, 2016). In VR, research has mainly focused on evaluating advanced graphics and capturing immersion levels and User Experience (UX) of interactions that stem out of the ‘typical’ and ‘expected’. The phenomenon Synaesthesia is of interest in the field of Human-Computer Interaction (HCI) as a way of creating new or enhanced experiences and interactions with MR technologies. This paper investigates how multimodal design characteristics can lay the foundations to a more ‘synaesthetic’ design approach in MR to identify how ‘atypical’ interactions can also affect UX in such environments.

The study was run remotely and 20 participants navigated around a series of mazes with puzzles. Emotion, immersion and presence were measured by surveys taken before and after the study. This was followed by a semi-structured interview conducted over videoconferencing about their experiences.

2. RELATED WORK

The synaesthetic-oriented approach to MR technologies is an underexplored area in HCI. The approach originates from the phenomenon Synaesthesia where people naturally combine senses resulting in, for example, being able to ‘see’ music or ‘hear’ colours among other sense combinations (Merter, 2017). The synaesthetic approach itself is a framework which combines sensory elements (Merter, 2017). Jaimes and Sebe (2005) found that multisensory VR research rarely combines the senses simultaneously. Diesendruck et al. (2010) used VR as a verification method to compare a synaesthete’s month-space perception to a control group. Our study aims to combine the

visual and audio aspects to create a synaesthetic-oriented UX that will inform a design framework for mixed and hybrid experiences in VR/MR.

In an attempt to design novel and gameful experiences in VE/VR/MR, existing in-person escape rooms have provided inspiration. Escape rooms are a series of puzzles solved by players in order to complete set tasks (Nicholson, 2016). Virtual escape rooms are used to perform tasks via different navigation routes and often within certain time constraints. In recent years, multisensory VR escape rooms have been established with the Hyper Reality Experience (2017) and The VOID (2019). Such attractions allow players to physically interact with a VE which has been mapped onto a real location. The player's senses are engaged by features such as a haptic feedback vest and a scent dispenser which activates at certain sections of the narrative (AWE – Augmented World Expo, 2016). Such triggers require the co-location of participants to experience these.

While there is limited research in VR for explicitly embedding 'novel' multimodal approaches, there is a good body of research that acknowledges the value of multisensory design in interventions that support healthcare and promote wellbeing. Multisensory VEs have been used to reduce stress levels (Putrino et al., 2020), have shown potential for improved quality of life for people with dementia (Cheng, Baker & Dursun, 2019; Sánchez et al., 2013) and provided interventions for anxiety symptoms (Rajasekaran et al., 2011) and the profoundly disabled (Brooks, 2021). Although multisensory design has been embedded in VR/MR there is limited research focus on embedding or facilitating 'fused' sensory experiences, for example, synaesthetic approaches.

Synaesthesia has been used to improve creative ideation by making cards featuring sensory elements which can be combined to inspire novel HCI designs (Lee et al., 2019). The combination of touch and motion has been used to improve possibilities of interaction with mobile devices (Hinckley & Song, 2011). The association between mood and music as well as colour and music has been used to create a new form of music player (Voong & Beale, 2007).

Fire training simulation has been a significant area of VE research. Heaters and a smoke smell which both increase in intensity when approaching the fires in VR have been used to replicate the feeling of being in a burning building (Shaw et al., 2019) promoting enhanced immersion. Shaw et al. (2019) and Smith and Trenholme (2009) have participants exhibiting responses that would not map to responses in a real world fire scenario such as opening doors with smoke coming from underneath them. Smith and Trenholme (2009) also show that training is required so there is not a discrepancy in

experience between those who play videogames often and those who do not when in VEs.

Multimodal interaction is another underexplored area of HCI where the multisensory elements are not 'fused'. Schifferstein (2011) created a framework for the design of multisensory experiences which we aim to expand upon in regard to MR and VEs. Colour-speech synaesthesia is based on multi-sensory perception (Bargary et al., 2009) and mirror-touch synaesthesia can affect people's perception of themselves (Maister, Banissy & Tsakiris, 2013). Multisensory perception has inspired art exhibitions (Casini, 2017) and cuisine (Spence & Youssef, 2019). It was hypothesised that a level of synaesthetic response is present in everybody (Casini, 2017; Spector & Maurer, 2013).

Perceptual phenomena such as change blindness have been used in VEs to redirect the user into taking a different path without realising it. For example, in Suma et al.'s study (2011) a participant can enter a virtual room, complete a task in it and proceed to the next room in a corridor. However, in reality, the user is entering and exiting the same room repeatedly as the location of the door is changing while they are occupied with the task. Out of the 71 participants, only one noticed the change and that was when prompted by the researcher. We aim to investigate change blindness within the context of synaesthetic-oriented approaches to see, for example, whether 'fused' UX retains this phenomenon. Are VEs that facilitate 'fused' UX more or less immersive? Do they sustain or promote change blindness? The investigation of such aspects would be valuable for applications in healthcare (supporting/training people with impairment), in transport (designing applications that assist drivers) and in fire and rescue services. Our aim is to see how the synaesthetic-oriented and multimodal approach can be implemented in Virtual Worlds (VEs and VR) and if they result in a new or enhanced (i.e. more 'fused') UX.

3. METHODS

3.1. Participants

20 participants were recruited using the snowballing approach (Patrick, Pruchno & Rose, 1998) both within and outside the University of Leicester. As the synaesthetic approach should be accessible to all, it was decided to recruit a wide demographic. There were 11 men and 9 women who participated. The age range was 18-65 with an average age of 36. Out of 20 participants, seven had never played videogames while the other 13 had experience playing apps or computer games.

3.2. Research questions and hypotheses

The research questions for this study are as follows:

- (i) What effect do multimodal elements have on immersion and presence in a Virtual Environment?
- (ii) What effect does a multimodal approach to a Virtual Environment have on user performance?
- (iii) How does a multimodal approach contribute to perceived experience and levels of immersion in a Virtual Environment?

The first research question is addressed by a quantitative analysis of data whilst the third research question is assessed with qualitative data. Following Witmer and Singer (1998) and Berkman and Akan (2018), presence is defined as feeling present in a VE while immersion focuses on the stimuli creating a feeling of interaction and inclusion within the VE. As higher immersion levels result in higher presence levels, both concepts are being measured simultaneously in this study.

The hypotheses are as follows:

- (i) Higher levels of immersion in participants would lead to improved problem solving performances as well as faster navigation.
- (ii) Higher levels of immersion would lead to higher levels of change blindness.
- (iii) Participants who notice the changes would be more likely to have an enhanced UX in regard to maze narration.

For example, in the third hypothesis, if the participant notices the changes, they would be less likely to follow the narration's instructions.

3.3. Study design

The design was within-subjects so participants tested every condition with the same sensory elements. Puzzle room is abbreviated to PR.

Table 1: Independent variables

Independent variables	Levels	Related hypotheses
Colour of maze walls	White, blue or red	(i)
Narrator	Narrator or none	(ii), (iii)
Audio from windows	Birdsong or none	(i)
Location of stairs in regard to PR	None, before, after, before and after	(i)
Door in PR1	Left or right side of room	(ii)
Window view in PR3	Cloudy or sunny	(ii)
Wall colour in PR5	White or green	(ii)

Table 2: Dependent variables

Dependent variables	Measured by	Related hypotheses
Immersion and presence	Questionnaire	(i), (ii), (iii)
Emotions	Questionnaire	(i), (ii), (iii)
Performance	Maze and puzzle completion times	(i)
Participants' responses to multimodality	Interviews over videoconferencing	(i), (ii)

3.4. Study materials

The emotions survey combined the Visual Analogue Scale (VAS) (Riva et al., 2007) and the Positive and Negative Affect Schedule (PANAS) (Watson, Clark & Tellegen, 1988). VAS consisted of seven emotions which the participant had to rank on a scale of one to ten. The PANAS survey included ten positive and ten negative emotions which were ranked from one to five on a Likert scale. The immersion survey was the WAS Presence Questionnaire (Witmer & Singer, 1998) which measured immersive elements and overall immersion/presence (I/P) levels. The surveys were hosted online with Jisc Online Surveys (2020).

While I/P and emotion data was collected through surveys, a log file captured all performance data and all qualitative data came from participant interviews including levels of change blindness.

The VE was hosted online and was created using the 3D modelling software Blender version 2.83 (2020). The images rendered were connected using HTML and JavaScript. To ensure consistency, participants were requested to use Google Chrome on a Windows computer or laptop.

Figure 1 is the medium fidelity VE map. Puzzle rooms are represented in blue. Landmarks are represented by 'LM' and dead-ends are shown by asterisks. 'S' is the start and 'E' is the exit of each maze which are highlighted green.

Figure 1: Medium-fidelity map of the study

We examined 'typical' multimodality by utilising sensory inputs/outputs such as auditory and visual triggers, while people were asked to perform certain tasks in a VE in an attempt to monitor some 'baseline' figures regarding overall UX and I/P

levels and identify whether certain immersion and attention-related phenomena, for example, change blindness and time perception skewing, are present within such settings. We wished to identify whether certain types (audio/visual/tactile such as clicking buttons to perform tasks) triggered specific subjective perceptions (positive or negative). By understanding individual modalities then there can be a better insight of how/when to 'fuse' them to simulate more synaesthetic-oriented UX.

3.5. Study procedure

To begin, participants accessed a website with instructions for completing the mazes and puzzles, a link to the emotion survey and a map for the first maze. Each maze had a collection of landmarks to aid navigation although only the first maze's map was available to participants as it was the largest. To navigate around the maze the participant used either the arrow keys or directional buttons at the top of the screen. Another feature was a help button with instructions. Due to the implementation of the maze, it was not possible for the participant to change their view point.

A notable landmark in the first maze was the open window (Figure 2). As the participant approached the final area of that maze, birdsong would play and become louder as they reached the window.

Figure 2: *The open window in the first maze.*

After the first maze was completed, participants moved onto the first puzzle room (Figure 3). There were four buttons on the table (red, yellow, green and blue) which each had an associated musical note. A four button sequence was played which the participant repeated by pressing the buttons onscreen. They had to successfully repeat four sequences to progress. Each participant had the same order of puzzles. Between the second and third sequences, the room's background changed while the camera focused on the buttons. The door on the left side of the room moved to the right side to measure the participant's change blindness. After the puzzles were completed an animation played of walking upstairs to a door in a first-person view. This was included to add a vertical aspect to

the navigation although it only played between certain mazes and puzzle rooms so it didn't become repetitive to the participant.

The second and third mazes were similar in design to the first except the wall colours were red and blue respectively instead of white. The second puzzle room had no changes in the background whilst the third puzzle room had a window with a view of a field on a sunny day which changed to a cloudy day overlooking a cliff.

Figure 3: *The first puzzle room before the door moves.*

The fourth and fifth mazes had landmarks, white walls and included narration which directed the participants to the exit. Participants were not made aware of this feature before completing the study. Narrators changed between mazes. The fourth puzzle room had no background changes. In the fifth puzzle room, the wall colour turned from white to green as the participant solved the puzzles. The duration of the mazes and puzzles was approximately 20-30 minutes.

After the final set of puzzles the participant completed the emotion and I/P surveys as well as downloaded the log file. A short interview with a researcher followed over videoconferencing. The interview was informal and asked about the participant's experiences with the study.

3.6. Data analysis

Qualitative and quantitative data was collected from the participants. Audio recordings of the interviews were transcribed then thematically analysed as described by Braun and Clarke (2006). NVivo 12 was used to assist the annotation of transcripts and the coding of themes. The survey responses and completion times formed the quantitative data. The data was analysed using IBM SPSS Statistics 26. Shapiro-Wilks showed that non-parametric tests were required. The two emotions surveys (before and after the study) were analysed using the Wilcoxon signed-rank test. Correlation tests used Spearman's correlation coefficient and demographics were compared using the Kruskal-Wallis test.

4. RESULTS & DISCUSSION

The study's results have been grouped based on the hypotheses from section 3.2.

4.1. Demographics

No correlations were found between either age or gender and any levels of I/P or emotion. Compared to those who played videogames, those who didn't had a more positive response to how compelling the sense of moving around the VE was (p -value = 0.048). Non-videogame players felt more involved in the visual aspect (p -value = 0.006) and audio aspect (p -value = 0.026) of the VE. Their senses were also more engaged (p -value = 0.019).

4.2. Hypothesis 1: Higher levels of immersion in participants would lead to improved problem solving performances as well as faster navigation

Results showed there was no correlation between I/P, problem solving and navigation. Instead, I/P was effected by positive and negative UX. Results showed emotions provoked by problem solving and navigation affected participants' perception of their performance.

4.2.1. Positive UX

Thematic analysis identified the participants feeling a **sense of improvement** over time but not merely as a 'learning curve' phenomenon but also as a more enhanced 'comfort-like' feeling. The non-standard controls initially provided a challenge to participants' I/P however, by the end of the study, they felt more competent using the controls.

'I think I got better at it because I knew what I was doing. The second and third time, I thought "yeah, alright, I've got to come back a bit and then I can look and see what other doorways I've got". So I was learning. I was learning, you know, all the time about, well, don't just assume that because you've got to go forward to do these things. Look back a bit further and see what your environment actually is and then you can work out where you're going. So, yeah, it was a learning game really.' [P2]

'I quite like the last one that wasn't the vocal one, if that makes sense. On the basis that, by then, you're getting used to it, you know. Instead of the first one, you're thinking and then by the third one you're thinking "yeah, I've got the hang of this now." And you feel more comfortable with it.' [P4]

Colour effects seem to have affected participants' perceived UX and usability (for example, reducing feelings of monotony and promoting task completion), something that can be particularly interesting within the context of synaesthetic experiences – indeed, it is well researched that

colour synaesthesia can affect cognitive processes such as memory. For a recent meta-analysis on the field please see Ward, Field and Chin (2019). The different wall colours in the mazes were received positively especially by participants who had spent a long time in the first maze with its white walls; the red walls of the second maze were a relief as it showed they had succeeded in progressing. It was suggested that the VE should be more colourful.

'That was quite nice 'cause the first one was just white and I'd spent so long in that and got quite frustrated so having a variation of colour was quite- was quite welcome.' [P1]

'Change the colour. Change the colour of the stairs or change the colour of the environment' [P13]

'Give it more colour! [...] It needed more colour!' [P16]

Landmarks as sensory triggers were positively received by users, especially around navigation. Some participants also expressed enthusiasm for the current design and had suggestions for a more detailed VE including a fire training simulation.

'Yeah, so I feel like if the water jug or something, like the water machine made a bit of a noise like humming or something, just different bits of sort of break up different areas of the maze.' [P11]

'I would choose the first one 'cause there were many objects, it was quite good.' [P13]

'That would be good for the mazes as well if there was a fire and you got smoke.' [P16]

4.2.2. Negative UX

A negative theme was **disorientation and confusion**. Participants who reported being lost in one of the mazes experienced this the most, often due to one of the study's conditions having a lack of landmarks and repetitive brick walls. Participants also struggled navigating due to the inability to change the camera's viewpoint to look around corners.

'So you know the walls were like the plain grey and as I got- I'm a bit- it was quite easy to get lost I suppose.' [P5]

'So I think, yeah, the hardest bits were not, you know, not being able to see around corners as it were, not being able to see behind you' [P7]

'Or there's something of note whereas when you spin round and all you see is blank walls, it's very... yeah, disorientating? Yeah, just feels like you're lost' [P15]

Another theme was **frustration** often stemming from the participant **feeling lost** within one of the conditions, often exacerbated by a feeling of lack of

proficiency with the controls. Another form of frustration was the perceived amount of time spent in the mazes. The first maze in particular was seen as frustrating by participants due to its size.

'the frustration of ending up in a blind ending' [P1]

'I mean, there were definitely a couple of them where I was getting a bit frustrated. I was like "I don't know!" Like, "I'm sure I've been down here. (laughs) I don't know how to get out of this maze!"' [P5]

'Probably the first one 'cause I just spent so much time in it. And it's not a nice memory though (laughs).' [P15]

'Oh, the bloody first one. It took me ages to get out of it.' [P16]

4.2.3. Overall change in emotions

Comparing the reported levels of emotions before and after the study showed a statistically significant increase in levels of 'Surprise' (p-value = 0.016), 'Pride' (p-value = 0.033) and 'Inspired' (p-value = 0.034). In the interviews, there was a theme of surprise which was attributed to the unexpectedness of the birdsong, the narration in the fourth and fifth mazes and the participant being surprised by their ability to complete the study. There was also surprise when the participant found out about change blindness however this would not be shown statistically as the surveys were completed before the interview. The birdsong was regarded as a pleasant surprise as both an indicator of a world outside the maze and as an indication that the maze was almost complete.

'[...] then I found the window and I was like "Oh! I must be close to the exit now."' [P19]

'[...] then the birdsong things coming louder was a really nice "Oh! This is uplifting finally".' [P15]

Participants were not informed about the narration before the study and some were surprised by its inclusion as it provided the correct route through the mazes. One participant believed it to be a mistake that was left in the final study due to its unexpectedness.

'There was an issue actually, I know you're going to go into the questions, but when I did the last two mazes, you could hear you giving directions' [P2]

'the first one with the verbal instructions 'cause kind of [took me?] by surprise' [P3]

Participants who initially felt like they may not have the ability to complete the study easily expressed surprise as well as a level of pride and accomplishment afterwards.

'[...] it worked quite well for me so I had more of a sense of "(pleased) Oh! I've done it!" You know, of achievement so it does get your emotions going certainly.' [P18]

'I think I was worried I was going to be rubbish so I was like just trying to smash through it.' [P5]

As well as reported increases in 'Pride' and 'Inspired' for all participants, non-videogame players reported higher levels of Inspired afterwards than those who play videogames (p-value = 0.034) as well as an increase in positive emotions overall (p-value = 0.008). On the other hand, there was a decrease in reported levels of 'Interested' (p-value = 0.01) and 'Enthusiastic' (p-value = 0.03) for all participants. Participants who expressed a lack interest often compared the non-standard controls and repeated puzzles negatively to commercial videogames. The average change in statistically significant emotions was taken from the survey responses and can be seen in Figure 4.

Figure 4: Changes in emotion from the survey.

4.3. Hypothesis 2: Higher levels of immersion would lead to higher levels of change blindness

The results showed while there was no correlation between overall I/P and change blindness, there was correlation with individual I/P questions. Furthermore, sensory elements affected I/P positively.

4.3.1. Change blindness

No participants noticed the change in the door's location in the first puzzle room (Figure 5). Two participants noticed the change in the window's view in the third puzzle room but only when prompted (Figure 6). Two participants noticed the change in wall colour without prompting (Figure 7). Nine participants noticed the change in narrators and two more thought there was a change but hadn't realised there were two narrators.

There was a positive correlation between noticing the visual changes of the maze (the door, window and wall colour) and how well the participant felt they could survey the VE using vision (p value = 0.484, p-value = 0.031). On the other hand, there was a negative correlation between noticing the audio change (the narration) and feeling involved in

the visual aspects of the VE (p value = -0.572, p -value = 0.008). In regard to the narration, certain participants mentioned that they stopped observing their surroundings once the narration began.

'as soon as the voice came over at the top, I stopped using my eyes. I just really was going based on sound. I just kept clicking left, whatever the voice told me to do.' [P7]

'I was thinking about it afterwards and I did totally just blindly follow those instructions.' [P7]

'ignored what I was looking at entirely and just followed the instruction.' [P8]

'I wasn't really using vision' [P20]

Figure 5: The door change in the first puzzle room.

Figure 6: The window change in the third puzzle room.

Figure 7: The wall change in the fifth puzzle room.

4.3.2. Sensory elements

A positive correlation was found between the sum of positive emotions and how involving the participants found the visual aspects (p value = 0.514, p -value = 0.021). Participant interviews emphasised how the landmarks around the maze

both helped with navigation and made the VE more visually interesting.

'[...] 'cause I was using those to try to pinpoint my way around.' [P5]

'[...] all those different bits sort of just made the environment seem a much more, instead of just a fake sort of situation, it did actually give it some sort of life to it.' [P11]

4.3.2.1. Hearing the Song

There was also a positive correlation between the sum of positive emotions and the audio aspects being involving for the participant (p value = 0.5, p -value = 0.025). The audio aspects discussed were the birdsong, the narration and the sound during the puzzles. The birdsong was mentioned without prompting by 17 out of the 20 participants. Participants responded positively to the sensory element as illustrated in the quotes below. Only one participant had a negative response finding it 'quite loud and sudden' [P12] as the rest of the first maze had no audio cues.

'That was quite uplifting. That was very nice. (laughs)' [P1]

'[...] a feeling of almost, [inaudible] more real feeling, something outside the room rather than just the background of a computer screen.' [P8]

'It was really warming to the senses' [P15]

'I think it made me feel that I'm not so much in an enclosed space now. I can see out.' [P18]

The narration was found by participants to be memorable and responses differed depending on how challenging the participants had found the initial three mazes. Participants who had felt lost previously appreciated the additional help whilst others preferred the challenge of navigating themselves now they had a good understanding of the format of the study and its controls.

4.3.2.2. Hearing the Colour

In the puzzle rooms, participants could choose whether they used the visual or audio cues or a combination of the two in order to remember the sequences. 13 out of 20 used only the colours either by visualising them, remembering the words or making the words into an acronym. One participant wrote the sequences down. Six participants used a combination of visual and audio aspects with one participant even stating that they had begun to associate the musical notes to their respective colours in a synaesthetic-oriented manner.

'It was like colours but kind of like it's the colour I'd see in my head. And then I think, by the end, I'd gotten used to the sounds that were like

associated with it so I was concentrating less on it and I could just like remember it.' [P20]

Later in the interview, due to the participant's interest in immersion, Synaesthesia (as a concept) was discussed and they expressed their surprise as they recognised that they had synaesthetic-oriented responses in the past.

'That's so weird, oh, my gosh! Yeah 'cause I think definitely, with different instruments, I associate different colours. Like with pianos, I probably would associate like darker colours just 'cause like how the colours of a piano usually is, whereas a guitar is more like colourful maybe? That's really interesting.' [P20]

4.4. Hypothesis 3: Participants who notice the changes would be more likely to have an enhanced UX in regard to maze narration

There was no correlation between noticing the changes and whether the participant followed the narrator's instructions. Instead, experience of videogames and user expectations affected the choice to follow the narration. Perception of UX was effected by realism as well as perceived passage of time.

4.4.1 Perception

Prevalent themes were participants attempting to second-guess the purpose of the study and the VE being a study rather than a commercial videogame effecting their perception. For example, some participants trusted the narration because it was a study rather than a videogame. While changes were being made visually in the puzzle rooms, before being told about this, a few participants thought the sounds were being changed or correct sequences were being rejected.

'I wonder whether- I'm not sure if you were deliberately changing the sounds a bit.' [P4]

'I did the, or at least the first time, I did the typical videogame thing of "don't go where the person's telling you" and, like, took a few wrong turns deliberately but, yeah, obviously then, uh, got back and was correct. I was just- I was constantly thinking, well again like too much playing videogames and puzzle games, just expecting something to be thrown in there to get in the way.' [P6]

'Yes, that's something I'd do in a Dark Souls to be honest but not in this game. I was expecting the voices to be honest. [...] They don't usually just mislead you like at the final bit with no warning whatsoever. [...] If it was a proper game that I was like just downloaded off the Internet or just, if I got that game on my phone from an app, I'd probably be less inclined to believe it. But I had a feeling that you made it and I didn't think you'd intentionally fudge the results of this so.' [P12]

'I was also thinking that you said you can attempt it as many times as you want but I was kind of conscious like: were you secretly monitoring that? But then the instructions said that that doesn't really matter so I disregarded that.' [P14]

'I thought "we're going upstairs. We're going higher. Is that relevant as we go higher up the building?"' [P16]

User expectations had an effect on their response to the study in regard to a 'feel for more' as well as expectations on how realistic or simplistic the VE design should be. Participants sometimes felt like there was a world outside the maze which could be seen and may be accessible through the window. Some participants also wanted a plot to increase their I/P in the VE.

'[...] there was the kind of feeling that there was more to it than you saw.' [P8]

'Yeah, even if they're silly [inaudible] a simple plot's, it's just a plot in general's quite nice.' [P12]

'I forgot that the goal was the door 'cause that was the first maze room. So I just went to the window and [then/I?] expected like to go into the window but then I looked at the map and then I realised I had to go near the door.' [P14]

'Well, yeah, 'cause you don't know what to expect, do you, so you've only got vague instructions. I thought that you could perhaps jump out the window or something.' [P16]

Realism vs. Simplicity. A **realistic design** was seen by some participants as better and realistic elements increased their I/P. Breaks in realism, such as effectively teleporting between rooms, affected their I/P in a negative sense. On the other hand, some participants felt a simpler design worked well for the purposes of the study.

'It felt more like realistic 'cause you can hear what is normal like the outdoors.' [P3]

'It's nice to have something to introduce you to the- instead of just appearing at [a/the?] doorway or appearing somewhere you're going into the room. It makes it slightly more interesting like, you know, somewhere real rather than just appearing in a room with no explanation of why you can't go back out the door you just came in.' [P8]

'it was quite simple actually I suppose but it worked well.' [P5]

'I liked how simple it was to be honest 'cause it was easier to navigate something that's got less distractions.' [L12]

Lost in time. When remembering the mazes, a theme was time and how it affected which mazes were the most memorable. The first maze was remembered for how long it seemed while any mazes which surprised participants with how quick they seemed were also mentioned.

'I thought "I couldn't have done that that quick, could I?"' [P18]

'They were sort of memorable 'cause they're so quick.' [P8]

'It seemed like an eternity (laughs).' [P16]

'The first bit was- it took me ages' [P1]

4.4.2.1. Time perception

There was a positive correlation between the sum of positive emotions and the participant losing track of time (p value = 0.488, p -value = 0.029). At the end of the study, participants were asked how long they felt they had spent completing the study. The time estimated by participants was longer than the real time spent in the mazes (p -value: 0.001) with 18 out of 20 participants estimating this. There is a positive correlation between the perceived time and how well the participant felt they could examine objects from multiple viewpoints (p value = 0.53, p -value = 0.025).

Figure 8: Perceived time compared to real time elapsed.

5. GENERAL DISCUSSION

Our study suggests that visual and audio aspects had a positive effect on UX both statistically and within interviews. Although the multimodal approach was only present in the puzzle rooms, one participant out of 20 showed a synaesthetic-like response between the colours and associated musical notes similar to a natural synaesthete's association between vocal pitches and colours illustrated by Baron-Cohen, Wyke and Binnie (1987). While the participant showed some evidence of synaesthetic-oriented tendencies, they were unaware of this before taking part in the study and were surprised when the synaesthetic traits were recognised in their own behaviour. This is a promising area for future research as if puzzle sequence completed 25 times in total can create an unexpected synaesthetic-oriented response, there

is a potential for a stronger response with more repetitions or synaesthetic combinations other than visual-auditory. Also, it is important to note that designing synaesthetic-oriented approaches should be experienced by all who participate in the multisensory VEs, not just those who have Synaesthesia as suggested by Casini (2017). Multimodality can potentially support such novel design approaches for MR innovation because the sensory elements do not prerequisite the sensory 'fusion'. This allows the framework of Schifferstein (2011) to be expanded on further in the context of MR and VEs.

Contrary to the second hypothesis, there was no correlation between noticing changes and I/P. This may be because so few participants noticed the changes and the participants who didn't follow the narration were influenced either by how they would act in commercial videogames or believed it to be a distraction technique. There was also no correlation between the overall I/P scores and the changes being noticed. Despite this differing to the hypothesis, we believe this could be a positive result for using change blindness in VEs. In comparison to Suma et al. (2011), the VE in this study was relatively simple and was hosted on a website rather than being in VR. However, a maximum of two participants out of 20 noticed each change. The fact that participants didn't notice large changes in an environment with few landmarks and there was no correlation with I/P suggests that a complicated VE is not required to distract most participants. As long as the participant is focused on a task, it is likely they won't notice changes in front of them. This can be particularly useful in the design of VR training suites as it could potentially bring the cost of an application down if the environments implemented don't need to be complex to be effective.

The levels of focus shown are also illustrated with the narration in the fourth and fifth mazes. Multiple participants mentioned focusing on the audio to the extent that they ignored the visual aspect. This is likely why the change in audio was detected the most out of all the changes and why there was a negative correlation between noticing the audio and the visual aspects of the VE being involving. The percentage of people who did not notice the change in voices was 55 per cent. This result is supported by Vitevitch (2003) who had 42 per cent and then 57 per cent of participants reporting 'change deafness' over two studies where participants had to repeat words said by a voice and, for some participants, the voice changed part way through the list of words. Vitevitch also runs an additional study to check that the voices can be easily differentiated. The narrators in this study were both women in their 20s however some participants who noticed the different voices felt they had slightly different accents. In regard to

following the narration, there is already research into multisensory fire training VEs (Wareing et al., 2018; Shaw et al., 2019). Using a voice to indicate the exit could be a useful addition to similar VEs as the act of leaving a building during a fire could be mapped onto participants following the narration in this goal-directed scenario.

Participants often overestimated the time spent on the study. This is contrary to research by Sanders and Cairns (2010) who found that their maze game resulted in participants underestimating the time taken. Block and Zakay (1997) found that people generally underestimate the time taken to complete a task. It is currently unknown why participants in this study overestimated the time taken. A speculation could be that the sensory triggers utilised as 'landmarks' (colour, objects, audio) had indeed an effect on time perception – however, this would need to be further examined. One participant had suggested they had become better at measuring time during lockdown (the study took place during England's second lockdown) however this is only anecdotal and research into the COVID-19 pandemic and its relation to time perception is outside the range of this study. Planned future research will include time perception as a sense in an attempt to understand why this has occurred.

There was a noticeable discrepancy in I/P scores based on participants' experience of videogames. Participants who often played videogames compared the study to commercial games and had a more negative response to the study due to this. In future work, we plan to use standard controls (arrow keys or WASD for movement, mouse for the camera) in order to map to user expectations. A tutorial would be necessary to teach non-videogame players the controls before the main section of the study to bridge the videogame playing skill gap between participants.

5.1. Implications for more 'Fused' Design

Participants discussed 'unexpected' emotional aspects promoted by combining multisensory I/O. Future designs to enhance UX for Virtual Worlds could include mixing multisensory aspects in unexpected ways (more synaesthetic-oriented) to trigger 'positive' surprises. More research is necessary to elicit what new forms for fusing sensory I/O would be more effective and indeed usable for users of Virtual Worlds. Another design research direction is to consider how to best track and transform situation awareness for 'fused' sensory environments to minimise fatigue and support more synaesthesia-oriented experiences.

6. REFLECTIONS AND FUTURE WORK

COVID-19 impact. An online-based follow-up to this study has been planned. It aims to expand the

amount and combination of senses incorporated in this study's VE (visual, auditory, time perception).

Initially, a VR study planned for summer 2020 was postponed due to the COVID-19 pandemic. The VR study featured the combination of the kinaesthetic (movement) sense with visual, olfactory and time perception senses.

The VR study was adapted to fit an online setting highlighting the challenges that VR research experiences under such crisis situations, acknowledging the constraints in implementing more extended multimodality. The kinaesthetic sense was adapted from the user physically walking along virtual corridors to navigating around a virtual maze using arrow keys. Of course, this cannot simulate the dynamics of a physical and broader kinaesthetic perception and experience but the challenges of remote VR for such study designs did not offer many alternatives. This is something that requires further discussion within the HCI community. The visual and time perception senses were simple to adapt however the olfactory aspect was replaced by audio for practicality reasons. Memorisation puzzles remained in both studies.

A website-based study was chosen due to the limitations of remote use of VR technologies. The amount of VR owners is considerably smaller than the amount of people who own an Internet connected computer. Moreover, all participants in an at-home VR study would have videogame experience which may decrease the variety of feedback. This paper's study has shown familiarity with videogames affected participants' UX.

Designing an immersive, multimodal experience without the use of typical VR technologies was an additional challenge. Initially, it seemed that only visual and auditory aspects would be possible with an at-home VE experience. As the results of this study suggest there was an increased level of I/P with these sensory elements, the follow-up online study will incorporate additional senses with the participant receiving a package of sensory props.

An additional constraint was monitoring and supporting participants in a remote study compared to being in a VR lab. A researcher was always available by email and when technical problems occurred, they took longer to solve than being in-person due to not being able to view the participant's screen or participants sometimes lacked the technical vocabulary to explain the issue. A possibility was to observe participants over videoconferencing using screen-sharing. However, this may have effected how participants interacted with the study if they felt they were being observed or judged. Moreover, participants with older hardware may not have been able to screen-share while running the study.

7. REFERENCES

- AWE - Augmented World Expo (2016) *Curtis Hickman THE VOID: Creating The Illusion of Reality*. Available from: <https://www.youtube.com/watch?v=Ebwtq1HZJ2A> (6 May 2021).
- Bargary, G., Barnett, K.J., Mitchell, K.J. and Newell, F.N. (2009) Colored-speech synaesthesia is triggered by multisensory, not unisensory, perception. *Psychological Science*, 20(5), pp.529-533.
- Baron-Cohen, S., Wyke, M.A. and Binnie, C. (1987) Hearing words and seeing colours: an experimental investigation of a case of synaesthesia. *Perception*, 16(6), pp.761-767.
- Berkman, M.I. and Akan, E. (2018) Presence and Immersion in Virtual Reality. *Encyclopedia of Computer Graphics and Games*, pp. 1-10.
- Blender (2020) *blender.org - Home of the Blender project – Free and Open 3D Creation Software*. Available from: <https://www.blender.org/> (6 May 2021).
- Block, R.A. and Zakay, D. (1997) Prospective and retrospective duration judgments: A meta-analytic review. *Psychonomic bulletin & review*, 4(2), pp.184-197.
- Braun, V. and Clarke, V. (2006) Using thematic analysis in psychology. *Qualitative research in psychology*, 3(2), pp.77-101.
- Brooks, A.L. (2021) Interactive Multisensory VibroAcoustic Therapeutic Intervention (iMVATi). *Recent Advances in Technologies for Inclusive Well-Being*, 196, p.325.
- Casini, S. (2017) Synesthesia, transformation and synthesis: toward a multi-sensory pedagogy of the image. *The Senses and Society*, 12(1), pp.1-17.
- Cheng, C., Baker, G.B. and Dursun, S.M. (2019) Use of multisensory stimulation interventions in the treatment of major neurocognitive disorders. *Psychiatry and Clinical Psychopharmacology*, 29(4), pp.916-921.
- Diesendruck, L., Gertner, L., Botzer, L., Goldfarb, L., Karniel, A. and Henik, A. (2010) Months in space: Synaesthesia modulates attention and action. *Cognitive neuropsychology*, 27(8), pp.665-679.
- Hinckley, K. and Song, H. (2011, May) Sensor synaesthesia: touch in motion, and motion in touch. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (pp. 801-810).
- Hyper Reality Experience (2017) *Hyper Reality Experience | Tick Tock Unlock*. Available from: <http://hyperrealityexperience.com/> (6 May 2021).
- Jaimes, A. and Sebe, N. (2005, October) Multimodal human computer interaction: A survey. In: *International workshop on human-computer interaction* (pp. 1-15). Springer, Berlin, Heidelberg.
- Jisc Online surveys (2020) *Online surveys*. Available from: <https://www.onlinesurveys.ac.uk/> (Accessed 6 May 2021).
- Lee, C.H., Lockton, D., Stevens, J., Wang, S.J. and Ahn, S. (2019, May) Synaesthetic-Translation Tool: Synaesthesia as an Interactive Material for Ideation. In: *Extended Abstracts of the 2019 CHI Conference on Human Factors in Computing Systems* (pp. 1-6).
- Maister, L., Banissy, M.J. and Tsakiris, M. (2013) Mirror-touch synaesthesia changes representations of self-identity. *Neuropsychologia*, 51(5), pp.802-808.
- McLay, R., Ram, V., Murphy, J., Spira, J., Wood, D.P., Wiederhold, M.D., Wiederhold, B.K., Johnston, S. and Reeves, D. (2014) Effect of Virtual Reality PTSD Treatment on Mood and Neurocognitive Outcomes. *CyberPsychology, Behavior & Social Networking*, 17(7), pp. 439-446.
- Merter, S. (2017) Synesthetic approach in the design process for enhanced creativity and multisensory experiences. *The Design Journal*, 20(sup1), pp.S4519-S4528.
- Niantic (2016) *Pokémon GO*. Available from: <https://pokemongolive.com/en/> (6 May 2021).
- Nicholson, S. (2016) The State of Escape: Escape Room Design and Facilities. *Meaningful Play* Available from: <http://scottnicholson.com/pubs/stateofescape.pdf> (6 May 2021)
- Oculus (2019) *Oculus Rift S: VR Headset for VR-ready PCs | Oculus*. Available from: <https://www.oculus.com/rift-s/> (6 May 2021).
- Patrick, J.H., Pruchno, R.A. and Rose, M.S., 1998. Recruiting research participants: a comparison of the costs and effectiveness of five recruitment strategies. *The Gerontologist*, 38(3), pp.295-302.
- Putrino, D., Ripp, J., Herrera, J.E., Cortes, M., Kellner, C., Rizk, D. and Dams-O'Connor, K. (2020) Multisensory, Nature-Inspired Recharge Rooms Yield Short-Term Reductions in Perceived Stress Among Frontline Healthcare Workers. *Frontiers in Psychology*, 11, p.3213.
- Rajasekaran, S., Luteran, C., Qu, H. and Riley-Doucet, C., 2011, January. A portable autonomous multisensory intervention device

- (pamid) for early detection of anxiety and agitation in patients with cognitive impairments. In: *2011 Annual International Conference of the IEEE Engineering in Medicine and Biology Society* (pp. 4733-4736). IEEE.
- Riva, G., Mantovani, F., Capideville, C.S., Preziosa, A., Morganti, F., Villani, D., Gaggioli, A., Botella, C. and Alcañiz, M. (2007) Affective interactions using virtual reality: the link between presence and emotions. *CyberPsychology & Behavior*, 10(1), pp.45-56.
- Sánchez, A., Millán-Calenti, J.C., Lorenzo-López, L. and Maseda, A. (2013) Multisensory stimulation for people with dementia: a review of the literature. *American Journal of Alzheimer's Disease & Other Dementias*, 28(1), pp.7-14.
- Sanders, T. and Cairns, P. (2010) Time perception, immersion and music in videogames. *Proceedings of HCI 2010* 24, pp.160-167.
- Schifferstein, H.N. (2011, October) Multi sensory design. In: *Proceedings of the Second Conference on Creativity and Innovation in Design* (pp. 361-362).
- Shaw, E., Roper, T., Nilsson, T., Lawson, G., Cobb, S.V. and Miller, D. (2019, April) The Heat is On: Exploring User Behaviour in a Multisensory Virtual Environment for Fire Evacuation. In: *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems* (pp. 1-13). ACM.
- Smith, S.P. and Trenholme, D. (2009) Rapid prototyping a virtual fire drill environment using computer game technology. *Fire safety journal*, 44(4), pp.559-569.
- Spector, F. and Maurer, D. (2013) Synesthesia: A new approach to understanding the development of perception. *Psychology of Consciousness: Theory, Research, and Practice*, 1(S), pp. 108-129.
- Spence, C. and Youssef, J. (2019) Synaesthesia: The multisensory dining experience. *International Journal of Gastronomy and Food Science*, 18, p.100179.
- Striem-Amit, E., Guendelman, M. and Amedi, A. (2012) 'Visual' acuity of the congenitally blind using visual-to-auditory sensory substitution. *PLoS one*, 7(3), p.e33136.
- Suma, E.A., Clark, S., Krum, D., Finkelstein, S., Bolas, M. and Warte, Z. (2011, March). Leveraging change blindness for redirection in virtual environments. In: *2011 IEEE Virtual Reality Conference* (pp. 159-166). IEEE.
- The VOID (2019) *The VOID | A Virtual Reality Experience*. Available from: http://web.archive.org/web/20200304070516if_/https://www.thevoid.com/ (8 May 2021).
- Van Kerrebroeck, H., Brengman, M. and Willems, K. (2017) Escaping the crowd: An experimental study on the impact of a Virtual Reality experience in a shopping mall. *Computers in Human Behavior*, 77(1), pp. 437-450.
- Vitevitch, M.S. (2003) Change deafness: The inability to detect changes between two voices. *Journal of Experimental Psychology: Human Perception and Performance*, 29(2), p.333.
- Voong, M. and Beale, R. (2007, April) Music organisation using colour synaesthesia. In *CHI'07 extended abstracts on Human Factors in Computing Systems* (pp. 1869-1874).
- Ward, J., Field, A. P. and Chin, T. (2019). A meta-analysis of memory ability in synaesthesia. *Memory*, 27(9), pp. 1299-1312.
- Wareing, J., Lawson, G., Abdullah, C. and Roper, T. (2018, September) User Perception of Heat Source Location for a Multisensory Fire Training Simulation. In: *2018 10th Computer Science and Electronic Engineering (CEECE)* pp. 214-218. IEEE.
- Watson, D., Clark, L.A. and Tellegen, A. (1988) Development and validation of brief measures of positive and negative affect: the PANAS scales. *Journal of personality and social psychology*, 54(6), p.1063.
- Witmer, B.G. and Singer, M.J. (1998) Measuring presence in virtual environments: A presence questionnaire. *Presence*, 7(3), pp.225-240.