Deep Subjectivity and Empathy in Virtual Reality: A Case Study on the *Autism TMI Virtual Reality Experience*

Jonathan Weinel, Stuart Cunningham and Jennifer Pickles

This document is an uncorrected proof/author draft of a book chapter forthcoming in the book New Directions in Third Wave Human Computer Interaction (Springer 2018 https://www.springer.com/gb/book/9783319733555)

1. Introduction

Autism is a "lifelong development disability that affects how a person communicates with and relates to other people, and how they experience the world around them" (NAS, 2016f). Among the features of the condition, autistic people[1] may experience increased or reduced sensitivity to certain sensory stimuli, which can cause everyday situations and public spaces to become extremely stressful and hard to deal with. 'Too Much Information' (TMI) is an autism awareness campaign produced by The National Autistic Society (NAS; a UK-based autism charity), which aims to highlight this issue, and the challenges faced by autistic people in public spaces (NAS, 2016e).

As part of the TMI campaign, the *Autism TMI Virtual Reality Experience* is an application that includes a short virtual reality (VR) film, which seeks to convey the experience of someone on the autistic spectrum in a shopping centre (NAS, 2016c). The application aims to give an impression of the overwhelming sensory experience that environments such as this can induce for an autistic person, while also highlighting the negative reactions that an autistic person having a 'meltdown' or 'shutdown' can produce from the general public (NAS, 2016b). The application provides this impression through means of a first-person point-of-view (POV) narrative, which makes use of graphical and sound design techniques to convey the subjective experience of an autistic child. In doing so, the film aims to promote an awareness of autism, and foster a greater understanding from the public, in order to reduce the negative reactions that further confound such situations.

The *Autism TMI VR Experience* can be seen as part of an emerging field of work in audio-visual media and VR, which seeks to use first-person POV as a means to simulate consciousness states or generate empathic tools. For example, the 360-degrees film *Clouds Over Sidra* (Arora and Pousman) tells the story of Sidra, a 12-year old girl who has spent the last 18 months at the refugee camp Zaatari of Jordan. Other examples of work in this field include those discussed at *Virtual Futures Salon: Electronic Empathy* (Mason et al, 2016), such as *The Machine to Be Another* (BeAnotherLab, 2016): a VR experiment in which two participants swap perspectives while performing the same physical actions, leading to a sense of exchanged embodiment; or Jane Gauntlett's *In My Shoes* (Gauntlett, 2016), an arts project aiming to produce empathy by recreating life experiences. Electronic simulations of auditory hallucinations are also being developed for the purposes of treating schizophrenia (Craig et al, 2015); while in entertainment, simulations of altered states of consciousness produced by drug experiences or psychosis are being used in an increasing number of video games (Cunningham, Weinel and Picking, 2016). These works are united through the use of audio-visual media and technologies such as VR as a means to convey conscious states in the first-person, using forms of telepresence to place the user inside the sensorium of another. The *Autism TMI VR Experience* can be seen in the context of such work, and seems to be one of the first examples of VR being used to communicate the experience of autism; though other interactive systems do exist, such as Hermann, Yang and Nagai's *EmoSonics* (2016), which seeks to communicate the emotions of autistic people through electronic sound synthesis.

Applications such as the *Autism TMI VR Experience* may provide powerful forms of communication. Yet the success of such media rests significantly on the extent to which they can make use of graphics and sound to provide effective and meaningful forms of communication. In attempting to provide simulations of consciousness, designers must overcome challenges in the use of graphics, sound and interactivity to convey the conscious experience of the subject. Since the 1980s, discussions of realism in computer graphics have tended to focus on the photo-realistic accuracy of graphical representations that reproduce shading, shadows and lighting in ways that approximate patterns of light entering the eye (Amanatides, 1987). In more recent studies such as Borg, Paprocki and Madsen (2014), these features continue to be used as benchmarks when assessing realism. Graphics and sound are key features of a VR system that may contribute towards its immersive capabilities, by generating a virtual environment that engulfs the senses, leading to provide the user with a sense of 'presence' within it[2]. Other related work such as Slater (2009) have sought to expand upon this area, by isolating 'place illusion' (PI) as presence, or the feeling of 'being there'; and 'plausibility illusion' (Psi) for the belief that what is occurring is actually happening. However, it seems that reproducing incoming visual and auditory stimuli that a person would see and hear from an external environment, if they were 'in the shoes' of another person, may not necessarily be sufficient to meaningfully communicate the subjective experience of what it is actually like *to be* that person. For instance, with regards to autism, the perceptual experience of incoming sensory information is significantly different than it is for someone who is not autistic, and thus designers must find ways to communicate this. As we shall see in this chapter, this may involve the use of various audio-visual techniques to suggest selective attention, non-aural or non-visual sensory experience, or emotional states. In this way, designers can not only create an impression of receiving similar visual or auditory stimuli as another person in a given location, but also begin to relate aspects of subjective perception that such stimuli may invoke. Recent applications such as the *Autism TMI VR Experience* seek to do this, yet to date we are not aware of any studies that evaluate the user experience of these designs.

In order to further our understanding of how VR applications may communicate subjective experiences, this chapter presents a case study on the *Autism TMI VR Experience*. We begin with an introduction to autism, which explores how the condition affects sensory experience and may cause high levels of stress in certain environments. Following this, we explore how the *Autism TMI VR Experience* communicates such experiences. First, we provide an independent 'expert analysis' of the application, which describes the narrative, graphical and sound design techniques that are used in detail[3]. We then discuss a quantitative pilot study that was undertaken to evaluate the user experience of the *Autism TMI VR Experience.* This study explores the extent to which selective attention, sensory experience and emotion are communicated effectively by the video, comparing both the VR and 360-degrees video versions. Through this, we reveal how participants may gain a sense of what it is like to be a person on the autism spectrum, and the extent to which VR supports this when compared with a 360-degrees video viewed from a standard LCD computer screen. Building upon the outcomes of this study, we propose three main types of first-person POV simulation, including simulations of 'deep subjectivity' that take into account multi-modal perception. We propose that 'deep subjectivity' more effectively communicates the perceptual experience of what it is like *to be* another person, and leads to more effective empathic tools. Finally, we conclude that the principles explored in this chapter are applicable not only for autism, but also other forms of first-person POV simulation.

2. Autism and Sensory Experience

2.1 What is Autism?

Autism Spectrum Disorder (ASD) as defined by the *Diagnostic and Statistical Manual of Mental Disorders* (DSM-5) (American Psychiatric Association, 2013) is a developmental disorder. It is characterised by impairments in social communication and interaction, and restricted, repetitive behaviours, interests and activities, which includes sensory sensitivities. Autism is a lifelong condition and people diagnosed as autistic will have presented with difficulties in these areas since early childhood. The condition is being increasingly recognised, and it is now thought that more than 1 in 100 people may have autism (Brugha and Spiers, 2012). It is a spectrum disorder, and the presentation is highly individualised in each area of impairment. Among the features that autism presents, differences in experiencing and processing sensory stimuli have been consistently recognised and reported since the condition was first described (Kanner, 1943; Asperger, 1944). Indeed, some of the earliest research stated that this should be a diagnostic feature of autism (Creak, 1961; Wing, 1969), however sensory sensitivity was not included among the diagnostic criteria until the DSM-5 in 2013. Recently research has also considered emotional response and regulation in people on the autism spectrum (Mazefsky et al, 2013; Berkovits, Eisenhower and Blacher, 2017). Despite this being a key area, particularly when designing intervention practices for behavioural responses, this area of research is in its relative infancy and further work is needed to determine the link between emotional regulation and behavioural response in autism. Increased awareness of autism is important within society, as this may lead to improved diagnosis and support, and hence tools such as the *Autism TMI VR Experience* are provided in an attempt to highlight features of the condition to the general population.

2.2 Sensory Sensitivities

Traditionally research has focused on the impact of autism on seven senses: gustation (taste), olfaction (smell), tactility (touch), proprioception (body awareness), the vestibular system (balance), vision and hearing (Bogdashina, 2010). Through first-hand accounts from autistic authors (Williams, 1994; Grandin, 2009) it has become evident that these sensory differences are apparent from the early years and can result in an altered perception of an environment from that of a neuro-typical person (Bogdashina, 2016). Sensory differences include both hypersensitivity and hyposensitivity to certain stimuli. Hypersensitivity may result in strong adverse reactions to certain sensory information, for example, finding certain sounds, such as a clock ticking, too loud to bear. Alternatively, an individual may experience hyposensitivity to stimuli, leading to dampened responses or not attending to certain sensory information (Pellicano, 2013). These differences in sensory perception are not equivalent across individuals, and may also vary over time; in different environments; and due to increased or decreased stress levels.

Oscillations between hypersensitivity and hyposensitivity, across any or all of the senses, can have a distressing or overwhelming impact upon the individual and can lead to feelings of disorientation and confusion. The constant bombardment of environmental sensory stimuli can be difficult to process, and may cause ‘meltdowns’ or ‘shutdowns’. Such sensitivity may also cause the individual to exhibit sensory-seeking behaviours, often referred to as selfstimulating behaviours, or ‘stimming’. These may include hand-flapping, rocking and humming, among others (NAS, 2016d). In some cases, these behaviours can be harmful; for example, head banging to seek an increased tactile response can obviously have negative consequences. Additionally, these behaviours can cause consternation amongst the general public, which in turn can create stressful social situations. Meanwhile, sensory stimuli can also have a negative impact on an autistic person’s ability to focus (Baranek, 2002), and thus careful consideration of an environment must be taken to adapt it in accordance with the needs of an individual (Pfeiffer et al, 2005; Kinnealey et al, 2012; Guldberg, 2010).

In response to the need to design environments that address the needs of autistic people, the NAS’s SPELL (Structure, Positivity, Empathy, Low arousal and Links) framework emphasises the need to create spaces that encourage low arousal levels to mitigate the possibility of sensory overstimulation, which can result in meltdowns or shutdowns. In some autism specific support services, environments such as sensory rooms have been specifically designed to deprive certain senses and enhance others, in order to reduce the anxiety that can occur from experiencing too much sensory information. For instance, a sensory hub may be soundproofed but filled with lights that change colour and brightness; or they may be completely dark but contain strong smelling materials. These sensory hubs are adapted to suit individual requirements, based on an individual’s sensory profile. Recently some research has been carried out exploring the use of interactive audio-visual technologies to aid the design of these environments and create multi-sensory instruments (Capellen and Andersson, 2016).

3 An Analysis of the *Autism TMI Virtual Reality Experience*

In order to increase awareness of sensory sensitivities that may be experienced by autistic people within public spaces, the NAS has created the TMI awareness campaign. As part of this, the *Autism TMI VR Experience* was created in collaboration with the design company Happy Finish. The application aims to demonstrate to users how a fairly standard environmental setting may appear to a person who has hypersensitivities to certain stimuli. It also demonstrates how the effects of this can build over time when appropriate support or reasonable accommodations are not put into place. The design of the video is informed by first-hand accounts from autistic people and individuals affected by autism such as parents and carers, and demonstrates the cumulative effect of adverse sensory information and the resulting ‘meltdown’ or ‘shutdown’. The report the video was based on utilised collective responses from focus groups, online surveys, polls, and qualitative analysis of personal accounts. This research collated the views of over 7500 people and was conducted on behalf of the NAS by Breathe Research, YouGov and nfpSynergy (NAS, 2015). The video is available to view through both a mobile application for use with VR-cardboard; and as a 360-degrees YouTube video (NAS, 2016a), the latter of which received nearly 200,000 views in the first 3 months since its launch.

In order to consider the design of *Autism TMI VR Experience*, the following subsections give a detailed account of the structural design and narrative; visual design; and sound design techniques that are used. This provides our 'expert analysis' of how these can be understood to relate to the conscious experience of a person on the autistic spectrum, and the type of effect that the audience may experience. This appraisal was carried out independently from the design process, and utilises an approach of critical analysis to evaluate the salient features of the digital media artefact as an artistic composition.

3.1 Narrative and Structure

For the VR-cardboard version of the *Autism TMI VR Experience*, the menu page states:

> For autistic people, the world can be a scary place. Everyday sounds, lights, colours, they can all be too overwhelming. We invite you to experience this for yourself and enter a world of Too Much Information. #AutismTMI.

A menu then allows the user to either visit the NAS website, or to view the video. The 360 YouTube version does not include these menus, but goes straight into the video part, which is the same on both.

The narrative of the video begins in a shopping centre, as a woman purchases a ticket. The film takes a first-person POV from a child (Alex's) perspective, as he waits while the woman (who we may presume is his mother) operates a ticketing machine. While he waits, the film shows us various sensory inputs that are a product of the environment. First we notice the glare of flickering lights, background noise and the muzak of the shopping centre. Next, we hear footsteps as a woman passes by. As another woman passes, she appears to step in something sticky that is spilt on the floor, and wipes her foot for a few moments before moving on. The scene begins to intensify as more people walk past, including a man holding a large collection of bright balloons; while in the background two ladies appear to be testing out some perfume. Alex's breathing quickens as he begins to panic, while his mother tells him to calm down as she finishes operating the ticketing machine, which appears not to be functioning as intended. Alex begins to panic as more sensory events unfold, such as a woman cleaning the floor. His stressed mother attempts to calm him, but by this point Alex is overcome by the situation and unable even to hear her clearly, as the scene overwhelms him. His breathing intensifies and the scene turns to black as he experiences a 'meltdown' or 'shutdown'. In the next scene, as he begins to calm down, we see a view of his mother and the car park, as his mother tells him to keep breathing, that everything is fine and that it is time to go home.

In the closing of the film, Alex delivers the message "I'm not naughty, I'm just autistic, and I get too much information", while text displays the web address of the NAS, encouraging the viewer to learn more and sign up the to charity.

3.2 Visual Design

As noted, the visual design of *Autism TMI VR Experience* assumes a first-person POV. This is practically achieved using a head-mounted VR camera, which presents the field of vision available from Alex's eye-view. This allows the visual material to provide stereoscopic 3D, which includes a sense of depth; and allows 360-degrees head tracking, allowing the viewer to look around from Alex's viewpoint. However, the film not only presents a first-person POV from the location Alex is in, but also does so in a way that seeks to describe the subjective experience of Alex using various visual techniques. As Alex blinks, the screen flickers black, and various visual effects such as graphical filters are used to describe his experience of the environment, or highlight specific sensory elements which capture his attention. This technique of subjective, first-person POV, which includes blinking and the use of digital effects processes, is notably used elsewhere in feature films such as *Enter the Void* (Noé, 2009) to describe the subjective experiences and hallucinations of a protagonist. However, in this case, visual effects are utilised specifically in order to give an impression of an autistic person in a situation where too much sensory information is causing panic.

Considering the visual effects used to create this impression in more detail, brightness parameters are temporally boosted to create an oversaturated flickering image, which suggests the sensory overload caused by the lights (0:14)[4]. Meanwhile, a temporal blurring filter effect suggests the discomfort or jarring effect that the environment is causing Alex to experience (0:19). As a woman walks across the scene, we see bright ripples of light emerging from her footsteps, giving an impression of how prominent and powerful the sound of her heels are for Alex (0:33). As people walk past, such as the man with the balloons, a filter is used to create an echo of the image, lending an impression of disorientation (0:42). Following this, a cloud of purple smoke is emitted from the characters in the background, who seem to be testing some perfume; the purple particle effect providing a visual means through which to describe the overpowering smells that Alex may detect (0:48). Alex attracts an accusing glance from a young man passing by in a red coat (0:55). Throughout this section we can also see Alex playing with a dinosaur toy in his hands (visible by looking downwards in VR), reflecting the need for tactile sensory stimulation.

It is notable that some of these effects provide a synaesthetic visual impression of the sensory experience across the modalities, since light is used to highlight footsteps, and purple smoke to draw our attention to powerful experiences of scent. Some of these effects are provided using 'keyframing', a visual effects technique that enables the animation of graphical parameters. As the film progresses, keyframed effects such as these are used with increasing frequency and intensity, to give a visual impression of the rising sensory overload that Alex is experiencing. As the film continues to show Alex's increasing sensory overload, a keyframed filter is ultimately used to blur the visual field and cause a 'tunnel-vision' effect that gradually closes in upon him as he enters a 'meltdown' or 'shutdown' (1:02 - 1:27). This technique is suggestive of the sensory overload Alex is experiencing, where he is completely overwhelmed by 'too much information', and is unable to cope with any individual elements successfully; instead he effectively closes down. Notably in the 360-degrees presentations of *Autism TMI VR Experience,* this visual effect constrains the viewer's field-of-vision, as darkness closes in, ultimately plunging the audience into blackness in a manner that some may find scary.

In the final section (1:35), a transition effect returns the image to a normal view of a car park outside the shopping centre, which is unaltered by the various graphical filters. The absence of visual effects, taken in contrast with the previous scene of the shopping mall suggests a calmer environment, which is less overwhelming for Alex.

3.3 Sound Design

The soundtrack of *Autism TMI VR Experience* describes the subjective aural experience of Alex, with the additional use of music to metaphorically relate his mood. At the beginning of the film we hear the background noise of the shopping centre, the footsteps of a person off-screen, and the voice of Alex's mother. As the flickering lights cause a swell in brightness, we hear a whooshing sound (0:14), while the diegetic background muzak[5] of the shopping centre swells. We also hear the sounds of Alex's quickening breathing, and the rustling as he plays with the toy dinosaur that he holds in his hands. As the woman enters the scene, who steps in something that is spilt on the floor, we hear a prominent squelching sound (0:24). The sound of Alex's breathing continues to quicken indicating his increasing stress, while more synaesthetic whooshing sounds, footsteps, his mother's voice, and the swinging of a metal door are heard (0:40). As an alarm sound is heard (0:43), the noise levels rapidly intensify, as we hear a loud spraying sound that accompanies the use of perfume in the background, and a high frequency ringing sound. Further clattering is heard in the background and more squelching sounds as a woman begins to clean up the floor spillage, and Alex begins to gasp as he becomes overwhelmed (1:00).

At this stage, the dialogue from Alex's mother become muffled; an effect that is achieved with a low-pass filter (1:00). As she drops her keys, the sound is extended using a time-stretching effect, creating a sustained high frequency drone (1:05). During the phase where Alex's tunnel vision begins to close in, we continue to hear his panicked breathing, amidst the background cacophony of dissonant noise textures[6] and high frequency drones; the noise becoming louder as the amplitude of his mother's voice decreases (1:02 - 1:27). In this way, the audio track reflects the shift of attention from the discernible voice of his mother as she attempts to calm him, to the wall of noise that gives an impression of Alex's anxious mood and experience of sensory overload.

During the transition phase where the visual field goes dark, we hear only Alex's stressed breathing, which gradually subsides as the scene fades into the car park (1:37). The ambient sounds of the car park fade in, together with the voice of Alex's mother, creating a mix of amplitude between parts that seems relatively normal, reflecting Alex's experience of sound, which is no longer overwhelming, but has now become manageable. As end credits and message are displayed, we hear Alex's voice as the narrator, who delivers the message "I'm not naughty, I'm autistic, and I just get too much information".

4. A User Experience Study of the *Autism TMI Virtual Reality Experience*

4.1 Pilot Study Design and Methodology

A pilot study was conducted to investigate perceptions of the VR-cardboard and the 360 YouTube video versions. A convenience sample of twenty (n=20) participants were recruited from a University campus and consisted of undergraduate and postgraduate students, faculty staff and administrative staff. Of the 20 participants, 1 was female and 19 were male. Each participant was briefed and randomly assigned to either view the VR video (n_{VR}=9) or the 360 video (n_{360}=11). Both groups of participants wore headphones during the experiment.

Participants experiencing the 360 version of the video were sat at a desk in front of a computer (an Apple iMac with a built-in 21-inch display screen, wireless keyboard and wireless mouse), where the YouTube video was maximised to fill the screen. They were then briefly introduced to the 360 video system by one of the researchers, and given instruction of how to change their viewpoint in the video should they wish to do so.

Participants viewing the VR version of the video were asked to stand to provide them space to move freely and instructed that they were required to hold the headset to their eyes for the duration of the video. As with participants in the 360 group, they were briefed about the use of the VR-cardboard headset and the ability to move their head to change their viewpoint in the video.

An online, web-based, survey was used to collect participants' responses, which was completed immediately after watching the video. The total process took approximately 10 minutes per participant. The principal aims of the pilot study were to determine:

1. The efficacy of both videos in portraying Alex's subjective experiences to the viewer.
2. If there is a difference in the viewer perception of Alex's experience in the video between the VR and 360 formats.

The survey consisted of 9 questions relating to sensory features and 3 questions relating to their overall impressions of the video and their sense of relationship to Alex. These 12 questions were mandatory and presented the participant with a statement, to which they were asked to indicate their level of agreement using a Likert scale with the following points: *1. Strongly disagree; 2. Disagree; 3. Neither disagree or agree; 4 Agree;* and *5. Strongly Agree*. In addition, there were two open-ended questions at the end of the survey where participants could provide free text responses, which were optional.

The sensory aspects of the video section of the survey explored how effective participants found the video was at highlighting various aspects of Alex's subjective experience. The questions also correspond with the seven sensory aspects considered important in autism research: gustation (taste), olfaction (smell), tactility (touch), proprioception (body awareness), the vestibular system (balance), vision and hearing. In addition, since selective attention mechanisms towards sensory inputs are important, we look at attention. Also, as sensory experience leads Alex to have an emotional response of high arousal and distress, we also explore emotion specifically.

Considering the sensory aspects of the video, the following null and alternate hypotheses were investigated:

> H_0: There is no difference in viewer perception of sensory aspects in the video between the VR-cardboard and 360 YouTube formats.

H_1: There is a significant difference in viewer perception of sensory aspects in the video between the VR-cardboard and 360 YouTube formats.

Sensory concepts were explored with participants in both groups by asking for their level of agreement with the corresponding statement as shown in Table I.

Table I: Sensory Concepts and Corresponding Survey Statements

Sensory Concept	**Statement**
Attention	*The video highlighted to me what caught the boy (Alex's) attention*
Gustation	*The video highlighted to me what the boy (Alex) was tasting*
Olfaction	*The video highlighted to me what the boy (Alex) was smelling*
Tactility	*The video highlighted to me what the boy (Alex) was touching*
Proprioception	*The video highlighted to me the boy's (Alex's) awareness of his body*
Vestibular	*The video highlighted to me the boy's (Alex's) sense of physical balance*
Vision	*The video highlighted to me what the boy (Alex) was seeing*
Hearing	*The video highlighted to me what the boy (Alex) was hearing*
Emotion	*The video highlighted to me what the boy (Alex) was feeling*

In addition, we also seek to extract participants' overall impressions of the video, which explore whether the video makes the user feel as if they are in the same location as Alex (what it is like to be *somewhere*) and/or if they gain a sense of what it is like to *be* Alex (what it is like to be *someone*). The distinction between these two is important, since the video seeks to convey Alex's subjective experience of the shopping centre, beyond simply replicating the environment.

Regarding the overall impressions of the video and the participants' sense of relationship towards Alex, the following null and alternate hypotheses are investigated:

H_2: There is no difference in overall impressions of the video between the VR-cardboard and 360 YouTube formats.

H_3: There is a significant difference in overall impressions of the video between the VR-cardboard and 360 YouTube formats.

The overall impressions concepts were collected via the study, through asking participants to indicate their level of agreement with the statements shown in Table II.

Table II: Overall Impression Concepts and Corresponding Survey Statements

Impression Concept	Statement
Somewhere	*The video made me feel that I was in the shopping centre*
Someone	*I felt like I was the boy (Alex) in the video*
Empathy	*Overall, the video made me feel empathy towards the boy (Alex)*

Finally, at the conclusion of the Likert response questions, qualitative data was collected using open responses, to ask which features participants found most effective in conveying an experience of autism, or to allow them to express other comments about the video they watched. The two questions asked of participants were:

- *What features (if any) of the video were most effective in giving you an experience of Autism?*
- *Do you have any other thoughts, impressions or comments about the video?*

4.2 Results and Analysis

4.2.1 Sensory Concepts

The results obtained from the sensory concepts questions are presented using a polar plot, as shown in Figure 1. It is seen from this set of responses that there is a general consistency between the two types of video. *Emotion, hearing, vision*, and *attention* all receive a consensus of agreement responses, with the mean of each group's response being ≥ *4 (agree)*. *Proprioception* receives the closest to a neutral response, with the mean response from the VR-cardboard group being 3.22 and the 360 YouTube group being 3.00, followed by *vestibular*, which has a mean rating of 3.00 from the VR-cardboard group and 3.45 from the 360 YouTube group. *Olfaction* and *tactility* indicate responses towards the disagreement end of the Likert scale, with the respective means for the VR-cardboard group being 1.89 and 2.22, and the 360 YouTube being 2.27 and 1.91. The remaining sensory concept of *gustation* has the most consistent disagreement concept, with both ratings being < *2 (Disagree)*.

[insert Figure1]

Figure 1: Sensory Concepts between VR-cardboard and 360 YouTube Videos

A one-way MANOVA test showed that where was no statistically significant difference in sensory responses based on the type of video (VR or 360), F (9, 10) = 0.29, $p > .05$; Wilk's Λ = 0.793, partial $\eta^2 = .21$. From these results, we maintain the null hypothesis H_0: There is no difference in viewer perception of sensory aspects in the video between the VR-cardboard and 360 YouTube formats.

4.2.2 Overall Impression Concepts

These are investigated to determine which features came across to participants viewing the video. Results for these three questions are presented in the bar graph shown in Figure 2. Both of the experimental groups (360 YouTube and VR-cardboard) show strong levels of agreement with respect to the somewhere and empathy concepts, with mean responses across both groups being > *4 (Agree):* the *somewhere* mean responses were 4.22 for the VR-cardboard and 4.18 for the 360 YouTube; the *empathy* mean responses were 4.56 for the VR-cardboard and 4.55 for the 360 YouTube. The responses for the *someone* concept are weaker, means for VR-cardboard of 4.00 and 360 YouTube of 3.82, but are towards the *Agree* end of the scale.

[insert Figure2]

Figure 2: Overall Impression Concepts between the VR-cardboard and 360 YouTube Videos

A one-way MANOVA test showed that where was no statistically significant difference in the sense of being another person experienced, based on the type of video (VR or 360), F (3, 16) = 0.57, $p > .05$; Wilk's $\Lambda = 0.99$, partial $\eta^2 = .011$. From these results, we maintain the null hypothesis H_2: There is no difference in overall impressions of the video between the VR-cardboard and 360 YouTube formats.

4.2.3 Qualitative Responses

Finally, we consider the qualitative, open-ended feedback that was provided in the questionnaire[7]. In response to the query regarding features of the video that were effective in giving an experience of autism, participants referred to a variety of audio-visual features in relation to the communication of sensory experience. For instance, 11 of 18 responses commented on audio features, highlighting the breathing sound effects in particular, while 5 of 18 responses discussed visual features. The communication of emotions such as stress or anxiety was described by 5 of 18 responses. Meanwhile, in response to the invitation for other general comments, participants noted the possible application of the video in educational settings, technical challenges they faced in using the VR cardboard, and their own emotional reactions to the video.

5. Unpacking Subjectivity in First-Person POV Audio-Visual Media

The results of the user experience study suggest that the participants found the *Autism TMI VR Experience* provides an effective means through which to communicate the subjective experience of an autistic person, leading to a sense of empathy: both groups in the study indicated high levels of agreement with the *empathy* overall impression $\bar{x}_{VR}=4.56$ and $\bar{x}_{360}=4.55$. This is achieved through the use of a first-person POV representation. However, this does not only communicate Alex's incoming visual and aural sensory inputs through graphics and sound, since it also uses these to provide an impression of other multimodal senses and emotions: both groups in the study indicated high levels of agreement with the *attention* sensory concept, $\bar{x}_{VR}=4.00$ and $\bar{x}_{360}=4.09$, and the *emotion* sensory concept, $\bar{x}_{VR}=4.56$ and $\bar{x}_{360}=4.73$. Through this, the first-person POV simulation does not only show what it is like to be *somewhere else,* $\bar{x}_{VR}=4.22$ and $\bar{x}_{360}=4.18$, but also to be *someone else,* $\bar{x}_{VR}=4.00$ and $\bar{x}_{360}=3.82$. This is an important distinction, which we will now further explore by unpacking the notion of first-person POV representations, in order to reveal three possible design approaches. Of these three varieties of first-person POV, the *Autism TMI VR Experience* utilises the third, leading to simulations of 'deep subjectivity' that may yield improved levels of empathy.

5.1 Layers of Subjectivity

In order to unpack different forms of first-person POV, let us consider the simplified model of sensory perception shown in Figure 3. During visual and auditory perception, incoming sensory information is received via an external environment (stage 1). In accordance with Broadbent (1958) and Treisman's (1960) theories of attention, cognitive processes of attention filter or attenuate incoming sources of information (stage 2). Attended sources reach perceptual awareness and are passed for higher level processing; while unattended sources are attenuated, and receive more limited processing (stage 3). Research regarding the 'McGurk effect' (McGurk and MacDonald, 1976) and neuroimaging studies of the auditory cortex (Callan et al, 2003; Calvert et al, 1997) show that in the second stage auditory cortex, higher-level processing and the multimodal integration of visual and auditory information occurs. Similarly, visual and auditory perception may invoke memories and emotions. Here these associative, multi-modal processes, which involve other cognitive systems, are collectively summarised as 'multimodal integration' (stage 4).

[insert Figure3]

Figure 3. Simplified stages of perception, which suggest corresponding categories of first-person POV audio-visual media.

5.2 Varieties of First-Person POV

The four stages of the simplified model of audio-visual perception shown in Figure 3 can be used to suggest multiple approaches for the design of first-person POV media that represents subjective experience. In what follows, we identify three distinct categories of first-person

POV media that correspond with these perceptual stages, describing simulations that take into account progressively higher or 'deeper' levels of processing. The features of each of these varieties are summarised in Table III.

5.2.1 Environmental Reconstruction

First-person POV simulations that provide 'environmental reconstructions' replicate stimuli similar to that experienced by the subject, without taking account of subjective attention or perception. These use 'Incoming Sensory Information' (stage 1) as a basis for representation (Figure 3). For example, we may conceive of videos that use a head-mounted camera and microphones, to reproduce the incoming visual and auditory stimuli that a person may receive in a given situation. For audio-visual media, these simulations are usually limited to reproducing sensory information from the visual and auditory domains, although in principal artificial stimulation of olfactory (Spencer, 2006; Murray et al, 2016) or gustatory (Ranasinghe et al, 2012) senses could also be used. We see examples of these videos regularly in the form of the Internet videos produced by rock climbers and mountain bikers, who show us what they see and hear as they engage in their respective sports. These simulations do not necessarily show us which visual or auditory entities within the scene the subject is directing his or her attention towards, except perhaps through physical motion where the subject turns his or her head to face the object of attention.

5.2.2 Subjective Attention

First-person POV simulations that provide 'subjective attention' utilise visual or auditory techniques to reflect the perceptual mechanism of attention, by attenuating unattended sources. These simulations use 'Attention Filters' (stage 2) and 'Perceptual Awareness' (stage 3) as their basis (Figure 3), since they represent the sensory experience as it is perceived after attentional filtering. In sound this can be achieved by adjusting the amplitude or applying filters to unattended sources, to attenuate their levels and reduce their prominence in the mix. Conversely, attended sources can have their amplitude levels boosted, or frequency equalisation processes can be used to make them seem more clear or bright, so that they 'stand out' more clearly. We see this approach in aspects of the *Autism TMI VR Experience,* since certain elements, such as the sound of shoes squelching through the floor spillage, are given relatively higher amplitude levels. Conversely, towards the end of the film, the sound of Alex's mother's voice is processed with a low-pass filter, creating a muffled sound that reflects reduced awareness. Using techniques such as these, it is possible for simulations to take the perceptual mechanism of attention into account, thus providing simulations of attended sources.

5.2.3 Deep Subjectivity

First-person POV simulations that provide ‘deep subjectivity’ utilise the audio-visual medium to reflect not only the perceptual experience of visual and auditory information, but also aspects of perception that are non-aural or non-visual. As described by ‘Multimodal Perception’ (stage 4), higher level processing involves the integration of visual and auditory information, and invokes other cognitive systems such as memory (Atkinson and Shiffrin, 1971) and emotions (Russell, 1980) (Figure 3). Sensory experience of gustation, olfaction, tactility, proprioception, and the vestibular system are not primarily visual or auditory. Nonetheless, it is possible to represent these using techniques, such as sonic or graphical metaphors. Indeed, our ‘expert analysis’ noted that this is used in various aspects of the *Autism TMI VR Experience*, although our user study also revealed that participants did not strongly detect these features. Similarly, our ‘expert analysis’ suggested that the video might communicate a sense of emotion through features such as its narrative and the use of dissonant noise textures that sonify Alex’s sense of emotional anxiety. Since emotion was strongly recognised by the participants in both the quantitative results, ($\bar{x}_{VR}$ = 4.56 ; $\bar{x}_{360}$ = 4.73), and the qualitative results, these features seem to have been more effective. In general, these poetic or metaphorical techniques give an impression of other aspects of perception, emotions, tastes and smells.

Table III: Representational content included in three varieties of simulation

Representational Features	*Varieties of First-Person POV Simulation*		
	Environmental Reconstruction	Subjective Attention	Deep Subjectivity
Incoming external audio-visual information	✓	✓	✓
Attentional filters		✓	✓
Sonification/visualisation of other sensory modalities: gustation, olfaction, tactility, proprioception and the vestibular system			✓
Sonification/visualisation of emotion			✓

5.3 Discussion

By proposing three distinct categories of first-person POV simulation, we have unpacked the distinction between simulations that show what it is like to be *somewhere* and what it is like to be *someone*. Although often an 'environmental reconstruction' may show what it is like to be in the same location as a person, this does not always give a deeper sense of what the sensory experience of that person may be like, as this varies between individuals. As we have seen in this chapter, this becomes especially important when considering autistic people, as they have a markedly different sensory experience than people who are not autistic. We propose that simulations of 'subjective attention' and 'deep subjectivity' may be more effective as a means to convey such subjective experiences, and tentatively suggest that these may also lead to a greater sense of empathy for the audience, since they communicate more essential aspects of conscious experience.

6. Conclusion

This work continues to develop a research trajectory explored by the 'Affective Audio' research team, which investigates issues of design and evaluation, of computational systems that represent subjective emotions and altered states of consciousness (see also Weinel et al, 2014; Weinel, Cunningham and Griffiths, 2014; Cunningham and Weinel, 2016). In this context we find unique interest in the *Autism TMI VR Experience,* since the video utilises immersive VR technology and audio-visual design in order to communicate the subjective sensory experience and emotions of an autistic person. While the application is highly novel, it can also be seen as part of a broader emerging trend in the use of VR and 360-videos to convey subjective consciousness and provide empathic tools.

Through our 'expert analysis' of the *Autism TMI VR Experience* we sought to identify the specific design approaches that convey the subjective sensory experience of an autistic person. Following this, the results from the pilot user study began to evaluate the efficacy of these approaches. Whilst indicative, because of the sample size employed, the outcomes of the study suggest that the *Autism TMI VR Experience* is successful in conveying aspects of attention, vision, hearing and emotion, of an autistic person. Conversely the gustation, olfaction, tactility, proprioception and vestibular senses did not seem to be communicated effectively through the app, even though our analysis found attempts to incorporate some of these features through the design of certain elements in the video. With regards to the latter senses, it is notable that these features are less commonly explored in mass media, such as television, films, video games, radio; hence both the techniques for their representation and audience exposure to such materials may be less well developed.

Despite the difficulties in representing some types of sensory experience, the *Autism TMI VR Experience* nonetheless seemed to give the viewers' an overall impression of not only being *somewhere* else, but also being *someone* else; and participants reported a sense of empathy from this experience. In this chapter we have tentatively suggested that this sense of empathy may arise due to the representation of non-aural and non-visual features of sensory perception such as attention and emotion. Developing these concepts, we have proposed three varieties of first-person POV including that of 'deep subjectivity'. In further research it would be useful to compare all three varieties of simulation proposed using bespoke software, in order to explore the effectiveness of otherwise of each for generating empathy.

Among the outcomes of the study, it is also notable that our comparison of the VR and 360 YouTube versions of the *Autism TMI VR Experience* suggested no clear benefit through the use of VR-cardboard. This may be significant for the developers of such tools, as the use of VR technology may be an unnecessary additional expense. However, we acknowledge that VR-cardboard offers inferior quality in comparison with more expensive VR hardware solutions, which may offer higher resolutions and greater levels of immersion, leading to a more distinct advantage in terms of empathy levels. A future study should compare the responses of participants using the VR-cardboard, 360 YouTube, Oculus Rift, and HTC Vive devices, to determine if the platform and imaging resolution have an impact upon participants' responses. Furthermore, there may be other benefits for using VR-cardboard that are beyond the scope of our study here, such as the novelty of the VR experience, which may attract a larger audience to the video; and the branding opportunities that the cardboard provides.

To conclude, applications such as the *Autism TMI VR Experience* present a variety of problems for research, some of which we have begun to explore through the case study presented here, while others have been identified which require further work. Such applications may provide powerful new communicative tools that are useful for generating awareness of autism, but equally the concept of 'deep subjectivity' that we have explored here may also be transferred to other first-person POV forms of media, such as those encountered in computer games, serious games, simulations and immersive entertainment media.

Acknowledgements

The authors would like to acknowledge: the NAS and Happy Finish for creating the *Autism TMI Virtual Reality Experience*; the participants who took part in the pilot study; and the reviewers, who provided valuable insights and recommendations.

References

Amanatides J (1987) Realism in Computer Graphics: A Survey. IEEE Computer Graphics and Applications, 7(1):44-56. doi: 10.1109/MCG.1987.277026

American Psychiatric Association (2013) Diagnostic and statistical manual of mental disorders (5th ed.). American Psychiatric Association, Washington.

Arora G, Pousman B (2015) Clouds Over Sidra [video]. Within: Extraordinary Stories in Virtual Reality. http://with.in/watch/clouds-over-sidra/. Accessed 7th November 2017.

Asperger H (1944) Die 'autistichen psychopathen' im kindersalter. Archive fur psychiatrie und Nervenkrankheiten 117:76-136. Reprinted in: Frith, U (ed) Autism and Asperger Syndrome. Cambridge University Press, Cambridge, 1991.

Atkinson RC, Shiffrin RM (1971) The Control of Short-Term Memory. Scientific American, 225(2):82-90.

Baranek GT (2002) Efficacy of sensory and motor interventions for children with autism. Journal of Autism and Developmental Disorders (32)5:397-422. doi: 10.1023/A:1020541906063

BeAnotherLab (2016) The Machine To Be Another. http://www.themachinetobeanother.org/. Accessed 7th November 2017.

Berkovits L, Eisenhower A, Blacher J (2017) Emotion regulation in young children with autism spectrum disorders. Journal of Autism and Developmental Disorders, 47(1):68-79. doi: 10.1007/s10803-016-2922-2

Bogdashina O (2010) Autism and the edges of the known world: Sensitivities, Language and Constructed Reality. Jessica Kingsley Publishers, London.

Bogdashina O (2016) Sensory perceptual issues in autism and Asperger syndrome (2nd ed.). Jessica Kingsley Publishers, London.

Borg M, Paprocki M, Madsen C (2014) Perceptual Evaluation of Photo-realism in Real-time 3D Augmented Reality. In: Proceedings of the 9th International Conference on Computer

Graphics Theory and Applications (GRAPP-2014), pp.377-386. doi: 10.5220/0004683203770386

Broadbent D (1958) Perception and communication. Pergamon Press, New York.

Brugha T, Spiers N (2012) Estimating the prevalence of autism spectrum conditions in adults: Extending the 2007 adult psychiatric morbidity survey. UK Data Archive, Colchester, Essex. doi: 10.5255/UKDA-SN-7082-1

Callan DE, Jones JA, Munhall K, Callan AM, Kroos C, Vatikiotis-Bateson E. (2003) Neural processes underlying perceptual enhancement by visual speech gestures, NeuroReport, 14(17):2213-2218.

Calvert GA, Bullmore ET, Brammer MJ, Campbell R, Mcguire PK, Iversen SD, David AS (1997) Activation of Auditory Cortex During Silent Lipreading, Science 276(5312):593-596.

Capellen B, Andersson AP (2016) Health Improving Multi-Sensorial and Musical Environments. In: ACM Proceedings of the 10th Audio Mostly: A Conference on Interaction with Sound, Norrköping, Sweden, pp.61-68. doi: 10.1145/2986416.2986427.

Craig TKJ, Rus-Calafell M, Ward T, Fornells-Ambrojo M, McCrone P, Emsley R, Garety P (2015) The effects of an Audio Visual Assisted Therapy Aid for Refractory auditory hallucinations (AVATAR therapy): study protocol for a randomised controlled trial. Trials 16:349. doi: 10.1186/s13063-015-0888-6

Creak M (1961) Schizophrenic syndrome in childhood: Progress report of a working party. Cerebral Palsy Bulletin 3:501-504. doi: 10.1111/j.1469-8749.1961.tb10406.x

Cunningham S, Weinel J (2016) The Sound of the Smell (and taste) of my Shoes too: Mapping the Senses using Emotion as a Medium. In: ACM Proceedings of the 10th Audio Mostly: A Conference on Interaction with Sound, Norrköping, Sweden, pp.28-33. doi:10.1145/2986416.2986456.

Cunningham S, Weinel J, Picking R (2016) In-game Intoxication: Demonstrating the Evaluation of the Audio Experience of Games with a focus on Altered States of Consciousness. In: Garcia-Ruiz, M (ed) Games User Research: A Case Study Approach. CRC Press/Taylor & Francis, Boca Raton, FL, pp.97-118.

Gabrielsson A, Lindstrom E (2012) The Role of Structure in the Musical Expression of Emotions. In: PN Juslin and JA Sloboda (eds) Oxford Handbook of Music and Emotion: Theory, Research, Applications. Oxford University Press, Oxford.

Gauntlett, J (2016) In My Shoes Project. http://janegauntlett.com/inmyshoesproject/. Accessed 7th November 2017.

Grandin T (2009) Visual abilities and sensory differences in a person with autism. Biol Psychiatry, 65(1):15-16. doi:10.1016/j.biopsych.2008.11.005

Guldberg K (2010) Educating Children on the Autism Spectrum: Preconditions for Inclusion and Notions of "Best Autism Practice" in the Early Years. British Journal of Special Education, 37(4):168-174.

Hermann T, Yang J, Nagai Y (2016) EmoSonics – Interactive Sound Interfaces for the Externalization of Emotions. In: ACM Proceedings of the 10th Audio Mostly: A Conference on Interaction with Sound, Norrköping, Sweden, pp.61-68. doi: 10.1145/2986416.2986437

Kanner L (1943) Autistic disturbances of affective contact. Nervous Child, 2:217-250.

Kenny L, Hattersley C, Molins B, Buckley C, Povey C, Pellicano E (2015) Which terms should be used to describe autism? Perspectives from the UK autism community. Autism: The International Journal of Research and Practice, 2016, 20(4):442-462.

Kinnealey M, Pfeiffer B, Miller J, Roan C, Shoener R, Ellner ML (2012) Effect of classroom modification on attention and engagement of students with autism or dyspraxia. American Journal of Occupational Therapy, 66(5):511-519.

Mason LR, Cheren C, Gauntlett J, Shaw J, Pan X (2016) Virtual Futures Salon: Electronic Empathy. 26th July 2016, Lights of Soho, London, UK. http://www.virtualfutures.co.uk/event/vfsalon-electronicempathy/. Accessed 7th November 2017.

Mazefsky CA, Herrington J, Siegel M, Scarpa A, Maddox BB, Scahill L, White SW (2013) The role of emotion regulation in autism spectrum disorder. Journal of the American Academy of Child & Adolescent Psychiatry, 52(7):679–688. doi:10.1016/j.jaac.2013.05.006

McGurk H, MacDonald J (1976) Hearing lips and seeing voices. Nature, 263:747-748.

Murray N, Lee B, Qiao Y, Muntean GM (2016) Olfaction-Enhanced Multimedia: A Survey of Application Domains, Displays, and Research Challenges. ACM Computing Surveys (CSUR), 48(4):56.

NAS (2015). TMI Campaign Report. The National Autistic Society. http://www.autism.org.uk/get-involved/tmi/about/report.aspx. Accessed 7th November 2017.

NAS (2016a) Autism TMI Virtual Reality Experience. YouTube. https://www.youtube.com/watch?v=DgDR_gYk_a8. Accessed 7th November 2017.

NAS (2016b) Can you make it to the end? Behind the Scenes [video]. YouTube. https://www.youtube.com/watch?v=RnqR71AJVH4. Accessed 7th November 2017.

NAS (2016c) Experience Too Much Information in Virtual Reality. The National Autistic Society. http://www.autism.org.uk/get-involved/tmi/virtual.aspx. Accessed 7th November 2017.

NAS (2016d) SPELL. The National Autistic Society. http://www.autism.org.uk/about/strategies/spell.aspx. Accessed 7th November 2017.

NAS (2016e) Too Much Information. The National Autistic Society. http://www.autism.org.uk/get-involved/tmi.aspx. Accessed 7th November 2017.

NAS (2016f) What is autism? The National Autistic Society. http://www.autism.org.uk/about/what-is.aspx. Accessed 7th November 2017.

Noé G (2009) Enter the Void [film]. Fidélité Films (co-production), Wild Bunch (co-production), BUF (co-production).

Pellicano E (2013) Sensory symptoms in autism: a blooming buzzing confusion? Child Development Perspectives, 7(3):143-148. doi: 10.1111/cdep.12031

Pfeiffer B, Kinnealey M, Reed C, Herzberg G (2005) Sensory modulation and affective disorders in children and adolescents with Asperger's disorder. American Journal of Occupational Therapy, 59(3):335-345.

Ranasinghe N, Nakatsu R, Hideaki N, Gopalakrishnakone P (2012) Tongue Mounted Interface for Digitally Actuating the Sense of Taste. In: Proceedings of the 16th IEEE International Symposium on Wearable Computers (ISWC), pp.80-87. doi: 10.1109/ISWC.2012.16.

Russell J (1980) A Circumplex Model of Affect. Journal of Personality and Social Psychology, 39(6):1161–1178.

Sanchez-Vives MV, Slater M (2005) From Presence to Consciousness Through Virtual Reality. Nature Reviews Neuroscience 6:332-339. doi:10.1038/nrn1651

Slater, M. (2009) Place Illusion and Plausibility Can Lead to Realistic Behaviour in Immersive Virtual Environments. Phil. Trans. R. Soc. B, 364:3549–3557. doi:10.1098/rstb.2009.0138

Slater M, Wilbur S (1997) A Framework for Immersive Virtual Environments (FIVE): Speculations on the Role of Presence in Virtual Environments. Presence: Teleoperators and Virtual Environments, 6(6):603-616.

Sonnenschein D (2001) Sound Design: The Expressive Power of Music, Voice and Sound Effects in Cinema. Michael Wiese Productions, Studio City, CA, pp.152-153.

Spencer BS (2006) Incorporating the sense of smell into patient and haptic surgical simulators. IEEE Transactions on Information Technology in Biomedicine, 10(1):168-173.

Treisman AM (1960) Contextual cues in selective listening. Quarterly Journal of Experimental Psychology, 12:242-248, doi: 10.1080/17470216008416732.

Weinel J, Cunningham S, Griffiths D (2014) Sound through the rabbit hole: sound design based on reports of auditory hallucination. In: ACM Proceedings of the 9th Audio Mostly: A Conference on Interaction with Sound, Aalborg, Denmark. doi: 10.1145/2636879.2636883

Weinel J, Cunningham S, Griffiths D, Roberts S, Picking R (2014) Affective Audio. Leonardo Music Journal, 24:17-20.

Wing L (1969) The handicaps of autistic children: A comparative study. Journal of Child Psychology and Psychiatry, 10(1):1-40.

Williams D (1994) Somebody somewhere. Doubleday, New York.

[1] There is some debate regarding terminology to describe autism; see Kenny et al (2015).

[2] The definitions of 'presence' and 'immersion' used here follow those described by Slater and Wilbur (1997), where presence describes the feeling of 'being there' facilitated by digital media that has 'immersive' technical features. For a further discussion of presence in VR see also Sanchez-Vives and Slater (2005).

[3] Note that since our analysis is carried out independently from the design process, we provide a critical analysis of the artefact, but do not provide further commentary on the design rational, prototypes, or other aspects of its creation.

[4] For convenience, the time references used here and throughout relate to the YouTube version of *Autism TMI VR Experience* (NAS, 2016a).

[5] Film sound can be considered in terms of diegetic sounds that relate directly to the world presented within the story, and non-diegetic audio such as music that has no explicit basis within that world (Sonnenschein, 2001). Here 'diegetic background muzak' refers to the music of the shopping centre, in contrast with the non-diegetic music that is also used to reflect Alex's mood during the video.

[6] These 'noise textures' could be considered as a form of music, and provide a means through which to communicate emotion. For a further discussion of how dissonant musical features may represent unpleasant emotions, see (Gabrielsson and Lindstrom, 2012).

[7] In the interests of brevity, the description here is provided as a concise overview. In a future development of this study, a more in-depth qualitative analysis could be included utilizing formal methods of thematic analysis.