Assessing User Experience in A Virtual Reality Crowd Simulation.

Jacob Greenwood

A thesis submitted in partial fulfilment of the requirements of the University of Lincoln for the degree of MSc by Research

College of Science, School of Computer Science

November 2017

i Acknowledgments

Dr Patrick Dickinson –

I cannot fit into concise words how much Patrick has done for this work. The project would 100% not be where it is without his support, guidance and hours of hard work he put in. He truly went above and beyond for me and this research. I feel truly thankful to have had him as my supervisor over the past year.

Dr John Shearer and **Dr Kathrin Gerling** –

Both John and Kathrin have been a great support during the project. Providing new perspectives and insights into areas of study. Kathrin especially provided her invaluable perspective on how to run the study taking it to the next level.

IntLab –

The IntLab research team I was part of at the University of Lincoln are a phenomenal group of people who always had some useful insight and knowledge to share. Honourable mentions include **Dr Olivier Szymanezyk** and **Kieran Hicks.**

Mum and Dad –

I can't not thank my parents. Who, as always, support me in every way they can and drive me to succeed. They are a constant source of inspiration. I know of no one better than you two. Thank You.

Family and Friends –

And finally, my family and friends who have shaped me into who I am and are a constant force for good in my life.

ii Abstract

Agent-based crowd simulations are used for modelling building and space usage, allowing designers to explore hypothetical real-world scenarios, including extraordinary events such as evacuations. Existing work which engages Virtual Reality (VR) as a platform for crowd simulations has been primarily focussed on the validation of simulation models through observation; that is the use of embellishments to enhance a sense of immersion or constrained studies of proxemics. However, human participation in crowd simulations also has the potential to provide richer and more informative simulation outcomes. This issue has not yet been widely considered by researchers and warrants further study of user experience and behaviour.

This work examines VR crowd simulation through the lens of user experience and simulation outcomes. A task-based simulation scenario has been created in which a participant walks freely, and interacts with agents using the same social-force model which mediates agent-to-agent interactions. It examines and reports the effects of crowd density on both the users affective state and behaviour, also comparing it with that of simulated agents. The results gained from this study indicate a significant increase in negative affect with density, measured using a self-report scale, it also shows significant differences in some aspects of user behaviour, such as increased instinctive reactions during high-density situations. This work then discusses how the results relate to VR simulation design for mixed human-agent scenarios.

Table of Contents

Word Count: 23606

iii List of Figures and Tables

1 Introduction

Crowds of people are a common phenomenon, present in many contexts, such as sporting events and festivals, shopping centres and transportation hubs. Although they are part of everyday life and usually do not cause significant problems they can be dangerous. The mismanagement of these crowds can lead to accidents and injury; in extreme cases, unexpected events can cause panic or other behaviours which lead to injury and/or loss of life. Examples of this can be seen in tragedies such as the Hillsborough Disaster (BBC News, 1989) in which many lives were lost. The Taylor Report (Taylor, 1990, 47) quoted saying “The immediate cause of the gross overcrowding and hence the disaster was the failure, when gate C was opened, to cut off access to the central pens which were already overfull.” and “no attempt was made to control entry to individual pens numerically and there was no effective visual monitoring of crowd density”. For this reason, the effective management of crowds is desirable and requires training and planning for various scenarios and outcomes.

A common method of predicting and planning for crowd behaviour is to use computer-based simulations. Such simulations potentially allow planners to predict the outcomes from many different scenarios, using variable parameters and can assist the design of buildings or transportation systems. For example, to examine the outcome of disasters, such as fire, or help with the planning of large-scale events. Various algorithms have been developed for crowd simulation, and many commonly used methods are “agent-based”. That is, they model the physical behaviours and interactions of individual agents, in reconstructed representations of the spaces under consideration (e.g. a proposed building design), while those agents are trying to achieve certain goals such as evacuation. Large numbers of agents are desirable to accurately predict behaviours, but can also create a high computational load. As a result, it is commonplace for simulations to use relatively simple models of behaviour.

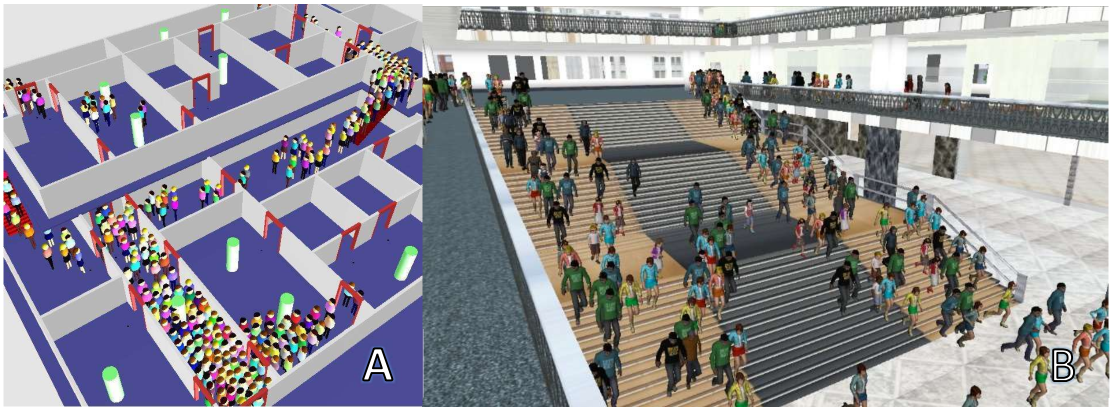


Figure 1: Examples of crowd simulations A: Pelechano et al (2007a). B: Jiang et al (2009).

One example of these simulation models, is the social forces model, which simplifies crowd behaviour to the agent's core motivations. These include reaching its desired destination, avoiding objects and other agents while walking there and sometimes wanting to walk with other agents or near different attractions (Helbing and Molnár, 1995). Within the simulation, these are represented by physical forces acting upon the agent. Research into crowd simulations has also examined the effectiveness and validation of different models (for example Golas et al, 2014 and Seitz et al, 2014) with the aim of, creating specialised simulations, increasing their believability and/or increasing the number of agents which can be simulated effectively.

Crowd simulations are usually designed to run with little to no input from the user at runtime, with the parameters and scenarios being set beforehand. Some models allow for limited real-time input, such as evacuation simulations which may allow the blocking of different locations to examine what would happen if a corridor became unusable due to a fire for example. This allows for the study of crowds in fixed scenarios and the effective animation of crowds such as those in movies like The Hobbit (Wired, 2015) but does not allow the user to interact with the simulated agents. Some video games such as Assassin's Creed: Unity (Ubisoft Montreal, 2014) and previous research (for example, Moussaïd et al (2016)) have allowed human-agent interaction in non-immersive environments.

This study will evaluate user experience when interacting with a crowd simulation in an immersive Virtual Reality (VR) system. This could provide new insights into the effectiveness of VR, the validity of simulation models and user response to varying crowd scenarios which would otherwise be problematic in the real world. More applications are presented in *Section 1.2.1.*

1.1 Virtual Reality

VR is a computer technology designed to generate realistic sound and images simulating the user's physical presence in a virtual world. Initially, this was difficult to achieve, requiring expensive bespoke systems such as Cave Automatic Virtual Environments (CAVE; *Section 2.6*) or custom head-mounted displays (HMDs). Recent technological advancements and commercialisation from large companies have created new accessible systems. These include the Oculus Rift and HTC Vive HMDs which allow for high fidelity tracking of the user's head and hands within a limited area. This along with a low latency 3D view of the world allows for an immersive experience.

1.1.1 Immersion and Presence

The sense of immersion provided by VR is the key element supporting the user's unconscious suspension of disbelief. This effect is called presence and allows the user to feel present within the virtual world causing their emotional and behavioural responses to be similar to that felt in the real world (Sanchez-Vives, Slater, 2005, 2). There are many ways to measure presence such as questionnaires, physiological measures and by observing the user's behaviour, see Whitmer and Singer (1998), Meehan (2001) and Wiederhold et al (1998) respectively. Meehan (2001) used both physiological and behavioural measures to track changes in user response between being in a standard sitting room environment and a "pit room" with a large drop in the centre of the room. He found that both measures indicated the participant had increased anxiety levels in the "pit room" within VR, even though the participant knew it was a virtual environment.

The hardware used is not the only factor for a successful immersive experience however. The design of the virtual world and characters are similarly important, and like other non-VR immersive experiences, the things that break immersion are not always obvious. For example, Slater's (2002) review of previous presence studies explain how virtual people with anomalies in their body shape broke the user's immersion less often than anomalies around their eyes and mouth. Other common breaks in immersion include graphical and behavioural anomalies, Pelechano et al (2008b) stated immersion was broken from characters jittering and colliding with each other and the world around them. This suggests that the virtual environment must faithfully reproduce reality sufficiently to avoid spoiling the experience.

1.1.2 Applications

VR provides many opportunities and benefits over conventional Human-Computer Interaction (HCI) methods, the first being video games. For example, The Lab (Valve Corp., 2016) and Star Trek: The Bridge Crew (Red Storm Entertainment, 2017) are both highly reviewed and popular among VR gamers, offering experiences not available in non-VR games. VR offers more uses outside of entertainment, such as in VR therapy (VRT; North, North, 2016; Opriş et al, 2015) offering a safe way for patients to experience immersion and exposure therapy while giving the physicians greater control over exactly what happens during the scenarios to better suit each patient. VR has also been used with training simulations to train pilots for example. This avoids the risks relating to inexperienced pilots flying real planes. The use of VR in research has been researched by many including Hupont et al (2015), they compared users' experiences of playing a training game with a 2D display to that of a VR alternative. They found that participants reported increased immersion and usability along with overall quality of experience (QoE). Ahlberg et al (2007) taught medical professionals using VR which showed increased ability post training, compared to those without VR training.

1.1.3 Limitations

Current VR systems are not perfect however, the lack of haptic feedback felt by the user restricts the possible uses compared to creating real-life scenarios. Other problems include the limited field of view, described as looking at the world through scuba goggles; Low image definition, due to the screens being closer to the user's eyes individual pixels become visible requiring much high definition than a conventional computer screen. All these issues can detract from the user's immersive experience potentially lowering the amount of presence felt, conflicting with VR's main objective. Nevertheless, contemporary systems still provide enough immersion allowing the user to feel present in the virtual world.

1.2 VR Crowd Simulation

This section will consider the possible applications of merging both VR and crowd simulations and outline possible limitations.

1.2.1 Applications

Using VR allows the user to see existing crowd simulations models from a new perspective, providing new insight into how realistic they are on the microscopic level. For example, Pelechano et al (2008b) used presence as a measure for determining the comparative realism of different simulation models. They suggest that unrealistic artefacts would diminish users' presence in a measurable way. It can not only allow for evaluation of existing models, the nature of VR helps a developer to test simulation models intuitively during development.

Applications for an effective and realistic VR crowd simulations could create a new method of training for emergency workers and event managers. Scenarios such as the evacuation of a burning building can be simulated with firefighter able to interact with the agents as if they were real, even with the given lack of physical interaction. People who fear crowds or activities involving crowds could have similar virtual therapy to other conditions, such as acrophobia (North and North, 2016). A validated crowd simulation in VR would allow research focused on the user response to different crowd scenarios, a subject that can be difficult to do due to lack of control over real life variables and difficulty measuring emotional response in real situations.

1.2.2 Limitations

Existing crowd simulations tend to focus on the general flow of the crowd, individual interactions are basic, with many models the interactions involve just walking past each other without gestures or other body language seen in real life. Researchers have aimed to improve this, for instance Narang et al (2016) created agents who would use gestures and make eye contact with the user, their findings indicate that participants prefer it to conventional simulations. Suggesting that these interactions can improve the user's sense of immersion. This is not easy to accomplish, human-human interactions are complex and difficult to simulate, and are currently outside the scope of modern AI.

1.3 Aims

This project investigates user response to a VR crowd simulation and corresponding behavioural artefacts. Little previous work has addressed this from the perspective of creating hybrid simulations, where the user acts an autonomous agent, interacting with other agents in the simulated environment. The specific aims are:

1. To investigate the affective response of human participants within a VR crowd simulation, particularly in relation to varying crowd density. Previous work has identified an effect of proximity of agents in VR, and I hypothesise that increased density will elicit a negative affective response.
2. To identify and quantify corresponding behavioural responses to varying crowd density, both in terms of reactive or gestural behaviours, and motion features such as walking trajectories.
3. To compare human behaviour in the simulation with that of other agents (undertaking the same objectives or task), as a baseline for human behavioural response, and discussion point for issues arising from the creation of hybrid human-agent simulations.
4. To identify, through qualitative analysis, any further issues relating to future development of mixed human-agent simulations, and help guide future work in this area.

1.4 Motivation and Contributions

This section will outline the contributions made and briefly summarise relevant results. It will then go on to present the initial motivations for this project including possible future research.

The experiment has shown a significant change in the affective state of participants when interacting with varying crowd densities. Specifically showing a significant increase in negative affect felt between density levels of $0.078m^{-2}$ and $0.425m^{-2}$. The relating change in behaviour has also been identified showing an increase in behaviours such as reactive behaviours and changes in trajectory to avoid agents

An Agent was created to carry out the same task as the participants, this allowed for comparisons to be made between human and agent behaviour. It was found that the agent spent more time closer to other agents than the participants did.

Interviews were held with each participant giving valuable insight into how the hybrid simulation felt when interacting with it. Behavioural aspects of the agents could be observed from a different perspective providing a possible improvement into how agents interact with other agents and people. For example, it was found agents very rarely stop for anything which some participants took as being “rude” and “aggressive”. This finding along with the human-agent behaviour comparisons suggest possible changes to be made to future simulation models

1.5 Structure of This Thesis

To begin, this thesis will review related academic literature for both crowd simulation and virtual reality, presenting their current uses and research recently being conducted into them. It will also review research being conducted into crowd simulations involving VR and how it has been used. Following this, *Section 3* will present the methods used to both develop and research the projects aim and then the implementation of the simulation with VR. *Section 5* details the evaluation that was conducted, including a description of the study, participants and the results with a findings section discussing these results forming links between them. Finally, a discussion of what the results could mean for future research and development with a conclusion for what this project has provided. The limitations of the study are also presented allowing future researchers to build upon this work.

2 Existing Work

This section provides a thorough review of current academic literature and outline areas of study this project will address. To begin, it will discuss crowd simulation in its current state and how it is used, the same is then done for VR specifically focusing on its role in human-computer interaction (HCI). The areas where these two fields have been combined will then be discussed, this includes simulation validation, user experience and finally crowd density and proxemics.

2.1 Crowd Simulation

As discussed in the introduction, crowd simulation is a field within crowd dynamics. It is the creation of simulated models to accurately portray the complex behaviour seen in densely crowded environments. This is not possible with conventional AI seen in most video game characters due to the sheer number of agents required and the complexity of $O(n^2)$ (Passos et al, 2008, 12). For this reason, the behaviour is often simplified to be a physics-based model with forces acting directly upon the agent pushing them passed other agents and obstacles (eg. Helbing and Molnár, 1995) or simple rule-based models (eg. Shao and Terzopoulos, 2005) for example. Zhou (et al, 2010, 6) categorises existing crowd simulation models into three groups:

1. Flow-based: this simulates the crowds as continuously moving flow of fluid. It allows for an extremely high number of agents to be simulated. It does not consider the actions of a single agent so is most useful for estimating the movement process of huge and dense crowds (Zhou et al, 2010, 6).
2. Entity-based: this model treats all agents the same, an example of this is Helbing's Social Force Model (Helbing, Molnár, 1998) with agents acting as particles influenced by forces representing 'various physical/social/psychological influences on an individual's movement in a crowd' Zhou (et al, 2010, 6).
3. Agent-based: Similar to Entity-based in that each agent is controlled individually, they are more intelligent however, with the ability to react and make individual decisions when faced with complex dynamic environments (Zhou et al, 2010, 6).
   - Zhou (et al, 2010, 6) goes on to say that the differences between entity and agent-based can be very small. For example, Braun (2003) uses Helbing (Helbing, Molnár, 1995) but then adds individual characteristics, such as an agent's likelihood to panic, making it both Entity- and Agent-based.

Although not always defined in this way there is evidence showing a clear difference between Flow-based and Entity-based models. Pelechano et.al. (2008a, 15) and Olivier et al (2014) defines them similarly with both macroscopic and microscopic systems linking to flow and entity/agent-based respectively (Zhou et al, 2010, 6). The requirements of the simulation dictate the ideal type of simulation. Large simulations, for example, may be too computationally intensive to simulate each agent, and smaller simulations may require greater fidelity to produce more realistic agent behaviour. Zhou et al (2010) have suggested the most likely uses for each type in *Figure 2*.

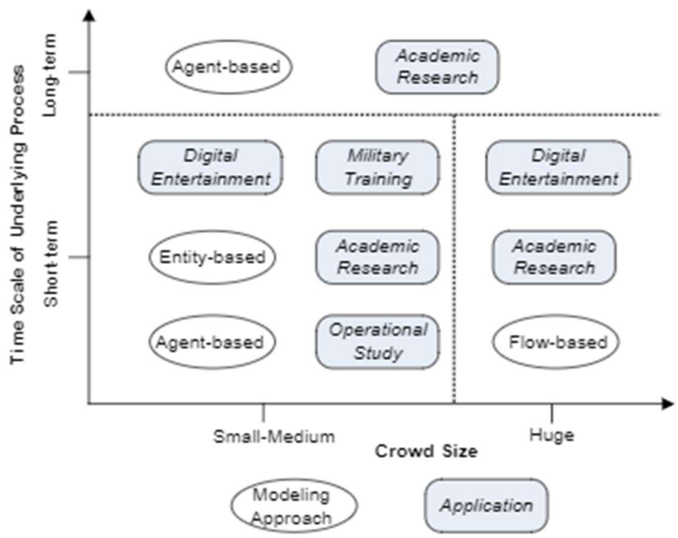


Figure 2: Classification of crowd models (Zhou, 2010, 5).

This project will require a small-medium simulation over a short period of time which from Zhou's figure above (*Figure 2*) suggests an entity or agent-based model. There are many different crowd simulation models across these two categories. One being the Cellular automata model in which the virtual world is split into many homogeneous segments in which an agent or a part of an agent can exist (Burstedde, 2001; Sarkar, 2000). Each segment then has a probability associated with it for the likelihood of an agent being able to move into it, this depends on the proximity of other agents and objects. This method allows for many agents to be simulated, its weaknesses, however, included the lack of continuous movement for the agent's due to the segmented world it also doesn't allow agents to make physical contact with one another (Pelechano et al, 2008b). Another type is the rule-based models, in which human movement is described through a set of basic rules, Reynolds' (1987) Boids are an early example of this. A third type and the model used for this project is the social force model, specifically the Helbing's empirical model (Helbing and Molnár 1995).

2.1.1 Helbing's Social Force Model

The social force model first described by Helbing and Molnár (1995) is built upon the idea that a pedestrian's motion within a crowd being predictable due to its habitual nature. It consists of three forces each representing a key motivation for the individual:

1. They want to reach their desired destination. Driving force.
    - This creates a driving force pulling them towards the next point on their path to the destination.
2. They want to keep a certain distance from other people/obstacles. Avoidance Forces.
    - A repulsive force pushing them away from people and obstacles
3. They may be attracted to other people or objects, such as friends or shop windows. Grouping Forces.
    - An attractive force pulling them towards select people/objects.

Other than the driving force these forces change depending on how far away they are from the entity, with avoidance forces increasing exponentially the closer they get. The speed and distance at which they increase are determined by the person's culture and personality which in turn affects the size of their personal space (Hall, 1968). An example of how this could be implemented is to simulate an agent in a rush, the driving force multiplier could be raised, and the avoidance multipliers lowered, causing the agent to collide with more agents but reach its desired destination quicker. From an observers' perspective, this agent may seem to be rude or careless. By doing the opposite, lowering driving force and raising avoidance, a nervous individual could be created.

A paper by Johansson et al (2008) aimed to parameterise the social force model by using video recordings of real crowds. By recreating video footage with the model, they could calibrate the forces until the results matched real-life observations. While doing this, they also further increased the validity of Helbing's model (Helbing and Molnár, 1995), specifically the exponential decrease in force strength over a distance was confirmed. They went on to try and improve the accuracy of the model, by using ellipses around the agents instead of the conventional circles. They found however that the small increase in accuracy was not enough to warrant the increased computational cost.

It is worth noting that the application of the social force model is not limited to crowd simulation, Mehran et al (2009) for example used it within computer vision to categorise real crowd behaviour. They found it could not only accurately detect abnormal behaviour but it also outperformed similar methods based on optical flow.

2.1.2 Advancements in Crowd Simulation

Research has shown that most crowd simulation papers focus on improving existing models or suggesting new ones. Improvements made to existing models include, Zanlungo et al (2011) who aimed to improve the social forces model by implementing collision prediction and avoidance; Durupinar et al (2008) who implemented the OCEAN personality model (Wiggins, 1996) into the HiDAC simulation created by Pelechano et al (2007a) allowed users to intuitively customise different agent's personality. Each of these after evaluation resulted in positive additions to their respective models. Rojas et al (2014) created a unique simulation focused on the group forming behaviour of pedestrians, they implemented a method by which the agents create varying group formations often seen in real life. Similarly Ahn et al (2012) developed a new model which re-used trajectories of real people, which allowed for small effects such as "zigzag hesitating movement" to be simulated which, they declared, cannot be formulated in traditional simulations.

Other papers focused on improving the computational performance such as Chen et al (2013) who created a hybrid simulation using both macro- and microscopic techniques depending on the situation. They managed to create a large (30000+) simulation with complex behaviours, compared to the standard (sequential) simulation they compared to could only achieve ≈8000. Passos et al (2008) aimed to make what they call a "Supermassive Crowd Simulation" by extracting the best performance available from their hardware. By creating a model using both the cellular automata crowd model (Sarkar, 2000) and Reynolds's (1987) boids they achieved a simulation with greater than a million entities, within interactive frame rates (>30fps).

Some papers are not directly creating or improving simulations but continue to study the behaviour of real people with the aim of furthering the field of crowd simulation. Seitz et al (2016) for example, states that most models focus on the pedestrian's torso behaviour even though it is directly dependent on their footsteps. They explain that the bipedal nature of humans is a large area of study in fields such as Biology and Medicine but has very little precedence in crowd simulation. Another study by Moussaïd et al (2016) used agents controlled by thirty-two participants in a single environment. Although not directly related to simulation it showed that behaviour was like that of real crowds.

2.2 Virtual Reality

VR is a term used to describe technology able to provide the user with the feeling of being physically present within a virtual environment. Until recent advancements this meant using bespoke technology that would often be big and intrusive, an example of this is a CAVE automatic virtual environment (CAVE; Neira et al 1993, 135). This involved projecting images onto the walls (including floor and ceiling if required) of a small room and the player standing in the centre as depicted in *Figure 3A*. Input would be achieved with a controller or other sensors depending on the system. These would be expensive to create requiring high definition projectors (Afanaan, 2017; Neira et al 1993, 141) and specialised technicians. Multiple studies discussed later will use this system. A more commercially available and user-friendly system can be found in head mounted displays (HMD; *Figure 3B*) however, these have seen a big increase in popularity in recent years due to better technology and commercialization from big companies such as HTC, PlayStation and Oculus. They display a 3D stereographic image and by tracking the headsets movement, the displayed image can be changed to correlate with the user's position, often paired with tracked controllers allowing 3D movement of both the users head and hands.

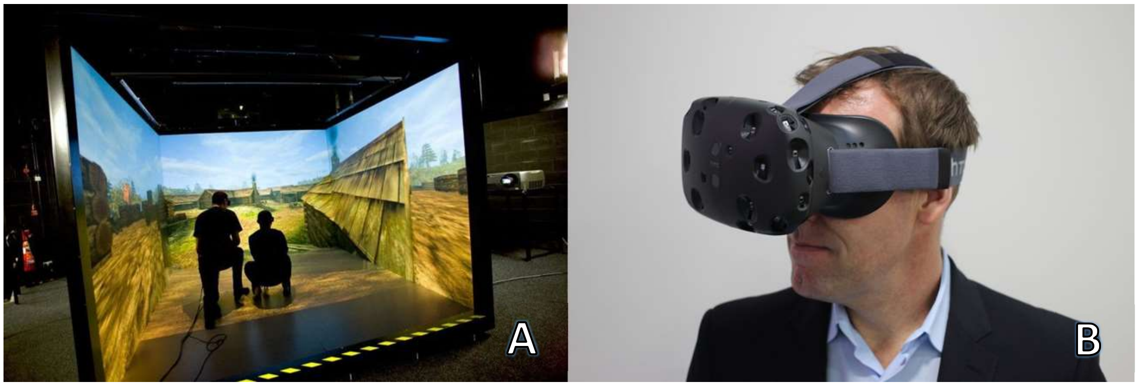


Figure 3: Two VR technologies. A) CAVE, displaying a virtual world projected onto walls and floor with the user sat in the centre (Afnaan, 2017). B) HMD, 3D stereographic display with acceleration, position and rotation tracking (Pesce, 2015).

As presented in the introduction VR offers many opportunities. North and North (2016) show that many studies have shown its effectiveness in therapy. Disorders such as agoraphobia, eating disorders and post-traumatic stress disorder were only a few of the conditions to be positively affected by VRT. The use in entertainment is clearly apparent in video games, but theme parks such as Alton Towers has created rides like Galactica utilising mobile VR. Other public VR experiences include museum exhibits; Jung et al (2016) studied the effects of VR and AR had on tourists and found greatly increased experience and increased intention of revisiting.

Other researchers have considered how they might improve current VR or make it more accessible, a paper by Laffont et al (2016) aimed to make the operation of VR devises for people with prescription eyeglasses easier. They state that the current method of adjusting each eye’s focus manually through trial and error is ineffective and my still leave them with non-corrected vision. To fix this they propose a system whereby the user can input their exact prescription and the headset will automatically adjust to facilitate them. Thomas et al (2016) created a system allowing multiple users to occupy the same space with little overhead in terms of price and deployment time.

Alaraj et al (2016) created a VR system to help train physicians in performing brain surgery, specifically, a method for real time haptic feedback for the simulation of aneurysm surgery. Results showed the they found it helpful in their education especially since the opportunity for training with this surgery are slim. Sugand et al (2015) found similar results with their “hip-screw” surgery simulator. These show that even though their haptic feedback systems were very discrete, they did improve the users experience and so as technology improves and more methods are theorised haptics will play a bigger role in VR. Other papers have suggested other haptic feedback methods for example Gugenheimer et al (2016) added gyroscopes to the headset to simulated kinaesthetic forces on the head, Achibet et al (2016) used passive elastics to simulate pushing objects, other full haptic suits are being developed by private companies such as the Teslasuit (2017). It is worth noting that no single technique for haptic feedback has gained precedence and is still an emergent technology.

An interesting paper by O'Brolcháin et al (2016) looked at the effects VR has on privacy when utilised in social networks. They state that the use of VR would allow companies to track the user's real-time response to the virtual world and people. Factors such as the user's emotional response could be recorded and analysed. They go onto present many impacts VR social networks could have on different forms on privacy and methods of combatting them in the future. Although this is unrelated to this project it shows the possible risks of using VR in the future.

2.2.1 User Experience and Presence

The effects VR has on users is still being investigated. Buttussi and Chittaro (2017) for example, recently compared the use of VR to 2D displays. The IPQ questionnaire (Schubert et al, 2001) was used to demonstrate a significant increase in the sense of presence whilst using a high-fidelity HMD, in an evacuation simulation. Similarly, a study by Hupont et al (2015) had participants play a serious game which aimed to teach the basics of forklift operation within VR. Using Witmer and Singer's (1998) questionnaire they found it achieved a higher presence compared to its 2D counterpart.

In the past, researchers have sought to use agents to either enhance user immersion or add additional naturalistic behavioural/ visual interactions. Garau et al (2005) for example demonstrated that user response is influenced by a range of factors, and to some degree, the user's pre-existing expectations for normal social behaviour mediated these human-agent interactions. Participants' experienced a high sense of personal contact when the agents responded to them. Narang et al (2016) similarly found that the users experience was improved when agents made eye contact achieving higher levels of presence.

Kyriakou et al (2015; 2016) studied how collision avoidance and other social interaction (gaze) affected the user's experience. By using a modified version of Witmer and Singer's (1998) questionnaire they found that collision avoidance increased the user's sense of realism, whilst other interactions increased their presence. Interestingly they also reported that the agents' active collision avoidance negatively impacted the users' performance but increases enjoyment. Similarly, a study by Sohre et al (2017) found that active collision avoidance increased the user's enjoyment.

A study by Egan et al (2016) aimed to measure the change in the QoE gained from using VR over non-VR systems in an objective way. By comparing the data from physiological sensors and a self-reported questionnaire, they found that the there was a correlation between them. They also found that the participants had an increased QoE whilst using VR, this result was limited, however, as only a small percentage of participants had used VR previously doubts were raised about whether the increased QoE was because VR is "better" or due to VR's potential novelty.

2.3 Crowd Simulation Validation With VR

The validity of a crowd simulation model is an important aspect and is an active area of research. This is because if a simulation does not accurately recreate real world events then its usefulness is severely limited. Historically this was done by comparing simulated crowd behaviour with that of real crowds in similar scenarios. Work such as Schadschneider and Seyfried's (2009) study compared the average speed of people and agents of increasing densities to create a fundamental diagram (Seyfried et al, 2008). Others such as Lemercier et al (2012) compared the formation of lanes and stop-go wave behaviour. These methods are limited however by the subjective nature and narrow view and do not evaluate how it would feel to be part of the crowd.

Pelechano et al (2008b) aimed to compare the validity of different models from an egocentric perspective, by using the user's sense of presence as a metric. They state that an effective simulation will produce a great amount of presence for the user. Therefore, any behaviour an agent presents that breaks this immersion is something to improve, making the simulation more valid. They give examples such as agents oscillating at high densities (common in social force models) and overlapping agents (models that do not use physics collisions). They used relatively small simulations however, with the independent variable being the simulation models itself. This project will build upon Pelechano's et al (2008b) work by looking at how density in a single simulation model will affect a participant's behaviour.

Mentioned earlier, Rojas and Yang (2013) and Rojas et al (2014) studied the group behaviour of agents within a VR simulation. They created a new method of simulating social groups and had participants interact with them in an immersive first-person way. Like Pelechano et al (2008b) they used the immersed user's instinctive "feel" for the crowd's realism. Their bespoke questionnaire, however, was not focused on immersion but directly related to the agents' realism and how natural the interactions felt. The findings were difficult to interpret however with no benchmark or comparisons were provided. Ahn et al (2012) also used immersive VR to evaluate their new collision avoidance model. They had participants stand stationary and watch a single participant, again a bespoke questionnaire was used evaluating what participants thought about the realism of the simulation. Recently Kim et al (2016) ran a study to validate their simulation and used both VR and 2D views. Results showed that users preferred the 2D top-down view, the task, however, was to compare the simulation with video source material, which was easier to do in 2D possibly contributing to the outcome.

The use of VR to validate simulations has become an active area of study, although work needs to be focussed on how to quantify the results into a clear comparable form. A summary and discussion of many VR crowd simulation studies has been done by Pelechano and Allbeck (2016). Apart from the previously mentioned study by Pelechano et al (2008b) however, the work offers less in relation to this project.

2.4 Crowd Density and Proxemics

Proxemics, first coined by Edward T. Hall (1966), is the study of spatial requirements of humans (and animals) when in groups and the effect increased density has on their behaviour, communication, and social interaction (Dictionary.com, 2017).

Figure 4 shows the four interpersonal zones; intimate, private, social and public along with their respective distances. With personal and intimate space reserved for interaction with close friends and family, social for conversations with associates and finally the public zone which is for speeches, lectures i.e. presenting to larger audiences. personal space, for example, is valued by most people and would cause negative emotions if it to be encroached upon by strangers, this helps explain why spending extended periods of time on crowded streets, buses, trains etc. can cause negative emotions. The distances presented below are not universal however and can vary with culture and individual personality.

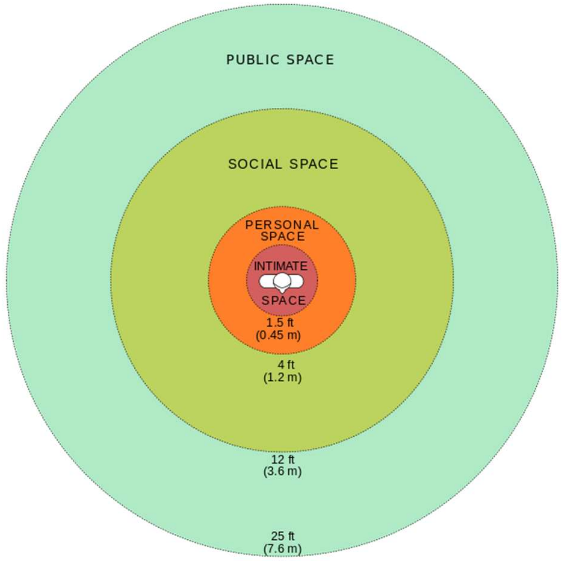


Figure 4. Edward T. Hall's interpersonal distances between people, radii are shown. (WebHamster, 2009)

The use of VR has allowed researchers to study the effect proximity has on users. Wilcox et al (2006) studied the effects of placing virtual characters and objects at varying distances from the participant (guided by the above values) to determine what effect VR had on their response. Physiological data and self-reported questionnaires showed similar emotional trends appeared between VR and real life.

Llobera et al (2010) used HMDs to determine what effect approaching characters had on the participant and found similar results to real-life situations. Interestingly, they also found that the characters form (inanimate cylinder vs human females) had little effect on response. Similarly, Christou et al (2015) had stationary participant approached by multiple agents who would stop at predefined distances or pass at different minimum distances. This also achieved similar trends to those seen in real life. These studies are limited however due to the participant having to stand stationary for the entire duration.

A study by Bruneau et al (2015) aimed to discover what factor make people decide to walk around groups of people or walk through them. Unlike Llobera et al (2010) and Christou et al (2015), these participants within a CAVE system could walk freely within the virtual world with a joystick. They found that the biggest influence on the participant trajectory was the time/ energy required to traverse the crowds. They then later used the knowledge to create their own energy based avoidance model.

These studies show the participant's instinctively respond to the agents' physical distance in a comparable way to that of real people. Other researchers, such as Kastanis and Slater (2012), have exploited this to manipulate participant behaviour, enabling them to be influenced in a predictable way. The above studies, however, did not allow the user to walk naturally through the simulation and were restricted to discrete scenarios.

3 Methodology

As discussed previous the aim of the project is to evaluate user experience in a VR crowd simulation. To do this an effect VR system will be used along with a crowd simulation model that is able to incorporate the user effectively and allow for them to interact with the world in an immersive manner. This section will detail the tools used to create and facilitate the study, including the chosen VR device and development environment. It will then go on the present the measures used during the study such as PANAS and interviews.

3.1 Toolsets and Machine Environments

A Discussion on which VR system is most effective is presented here along with its main benefits and how it will be used. The chosen development software is also presented.

3.1.1 Virtual Reality systems

There are multiple different types and models for VR systems, and although it was discussed earlier (*Section 2.2*) this section aims to outline the main contenders for which VR system will be used in this project and explain their pros and cons to finally come to a decisive conclusion on why the chosen system, HTC Vive, was ideal.

The gold standard for many years were CAVE systems, these could allow near 360˚ 3D experiences with no physical impact on the user other than having to stand in the centre of the room. Handheld controllers were often used for input allowing the virtual character to move freely in the virtual world. As mentioned earlier these are expensive and technically difficult to implement. They also do not offer as much versatility as other modern VR systems such as 1:1 head and hand tracking. This is also the case for mobile VR, such as the Samsung Gear, which tracks head movement but does not track hand or body movement. These also require a phone to run so any system developed would have to have lower computational costs which is a challenge for regular VR applications even without a full crowd simulation. For those reasons, CAVE and mobile were not a practical choice. This leaves two main options for VR the Oculus Rift and HTC Vive, they both offer 1:1 head and hand tracking allowing immersive interaction within a virtual world; such as, picking up objects and ducking around obstacles and agents. The HTC Vive is more suitable however due to its room scale capabilities allowing the user to walk around freely in a 4.6mx4.6m area compared to the

Rift's standing[1] only system. One issue with both the Occulus and Vive is the required wired connection it needs to a computer, this could potentially impede movement, possibly causing a trip hazard or reducing users' sense of immersion. Steps can be taken to avoid this such as suspending the wire above the user's head (Figure 5), more about this is discussed in *Section 5.1.2*.

Figure 5: Lab setup used to suspend the cable above the user's head, the project is near the centre of the play area used by participants.

3.1.2 Engine

Unreal Engine (Epic Games, 2015) was used to create the application for the study, this is because at the time it had the most support for VR systems. Pre-made assets such as a character controller were available allowing the developer to focus on aspects unique to this project such as

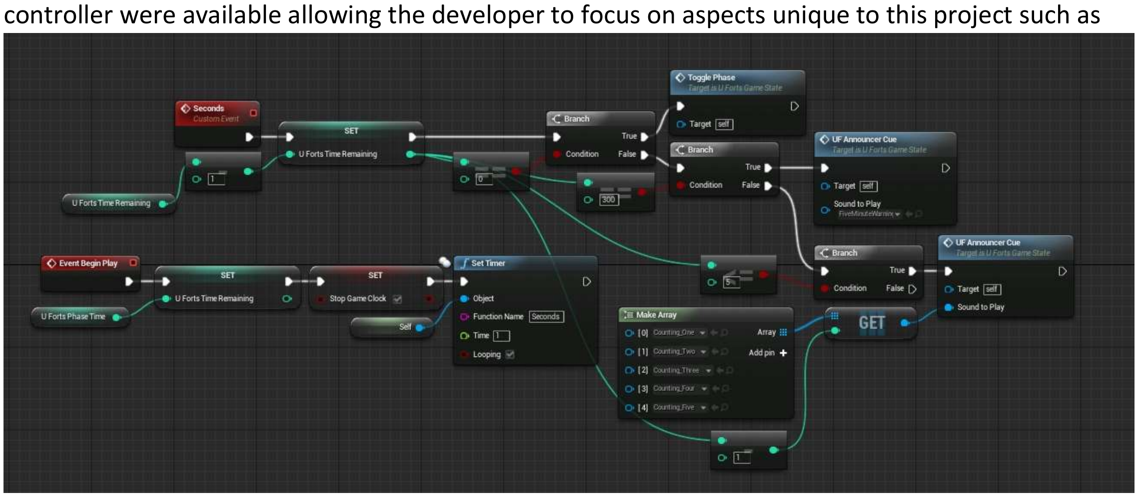


Figure 6: Unreal Engine's (Epic Games, 2015) visual scripting language, Blueprints.

[1] Although it is called standing only the user is free to move around in an area as long as they are close to and within field of view of a single camera. Allowing for side steps from the centre

the VR crowd simulation integration. Along with this Unreal has well-defined documentation and a large following allowing for great amounts of support from public forums and official sources.

Unlike other popular engines like Unity (Unity Technologies, 2016) for example, Unreal engine has a visual scripting language (shown in *Figure 6*) allowing quicker additions and changes compared to conventional languages such as C++ or C#. Other features include behaviour trees (finite state machine) allowing for quick and intuitive control over the agent's decision making and behaviour. More detail on the development can be found in *Section 4.*

3.2 Research Methods

One of the objectives is to assess the user's affective state during the simulation, which requires an appropriate measure. Emotional state is an inherently difficult aspect to quantify, however, so multiple dependent measures were used. A mixed research method of both quantitative and qualitative data was used to gather data, this played to the strengths of both (Morgan, 1998). Quantitative data supports the collection of large of amounts of measurable numerical data, that can be statistically analysed to provide objective deductions into the trends and irregularities that arise. Whereas qualitative data provides insight and explanation to help gain understanding of the narrative or inductive data collected during interviews with the participants. Thereby enabling a more comprehensive approach than might otherwise have been achieved. The section with present the three quantitative measures along with a detailed breakdown of the semi structured qualitative interviews.

The study will follow a within subjects' design, meaning each participant will experience every condition. This allows for comparisons to be made between scenarios with less participants than would be needed with a between subjects' design. One disadvantage to this however is the possible carryover effects experienced, ie previous conditions effecting the users experience during subsequent conditions. To counterbalance this the conditions will be presented to participant using a Latin square order, so every combination of orders is completed with the aim of negating its affect.

3.2.1 Positive and Negative Affect Schedule (PANAS)

A popular way to determine a user's emotional response is to use a physiological measure such as electro-dermal activity (EDA) to measure arousal, both Llobera et al (2010) and Christou et al (2015) used it in their VR studies considering the effects of proximity on the participants. Egan et al (2016) also used EDA and Heart rate and determined they can be used as a metric for QoE. An issue with physiological sensors, however is its unreliability during physical activity (Schumm, et al, 2008). Within these three studies the users were stationary, avoiding any such issues. This along with the difficulty to interpret in terms of affective response and user experience made it unsuitable for the project. Instead, a validated (Crawford, Henry, 2004) self-report measure was used, the Positive and Negative Affect Schedule (PANAS; Watson, et al, 1988). This is comprised of twenty words relating to both positive and negative emotions, the participants were asked to rate each one, one – five, depending on how they were feeling at the time. This allowed for the positive and negative emotions be analysed independently. PANAS has not been used for VR-based crowd simulation studies in the past but has been used in a similar area of study, games user research (see examples, Jennett et al, 2008; Russell and Newton, 2008). Based on previous work which suggests an increased discomfort due to crowding (Hall, 1965; Llobera et al, 2010; Christou et al, 2015) the adopted null hypothesis was: H_0= *There is no measurable difference in negative affect due to virtual reality crowd density.*

3.2.2 Trajectory and Observational Data

Trajectory Data was gathered to analyse the proximity of the user to agents while conducting the tasks. The speed and movement were also monitored. To do this the position and orientation of the user and agents were logged approximately 30 times a second along with other events such as items being picked up and task completion time/order. A tool was then developed to analyse this data. While this gives objective data into the user's behaviour in the system, it did not allow for their real-world behaviour to be observed. For this reason, observational data from video recordings (*Figure 7*) of the study were analysed. This allowed for the identification and quantification of common behaviour patterns between participants, and to further examine anecdotal accounts from previous works regarding participant behaviours (e.g. Pelechano et al, 2008a; Narang et al, 2016). The method used to facilitate the analysis of this data is explained in S*ection 3.2.2.1*.

3.2.2.1 Video Codes

Figure 7: Camera setup and view. Note the cameras ability to view the whole play area and what the participant is current seeing, through the projector.

The analysis of the video recordings required a method of categorising different behaviours observed enabling them to then be quantified and statistically analysed. To begin, the primary and secondary researchers made independent notes on everything that happened during each recording. This included time stamps and full-English descriptions of the users' behaviour*. Table 1* shows an excerpt from a single task.

Map Num	**1**
Map Name	**Medium**
Time	Description
0:28	task start, instantly notices agents and looks at them
0:31	waits for agent to pass and walks behind it
0:36	turns and looks around for any agents, begins walking before looking left
0:38	actively walks round back of agent as it crosses path
0:43	sees agent about to cross path, waits, looks left before continuing
0:46	walks behind agent
0:54	waits for 2 agents to cross in front before walking, moves hand to avoid touching agent
0:57	thinks got wire caught around arm pulls back and over as if wire is under arm. It wasn't.
1:04	looks both ways
1:06	walks behind and follows agent at same pace
1:15	looks both ways, quick steps to walk in front of agent
..	..

Table 1: Excerpt from initial observations with full descriptions of behaviour and time stamps.

This took a large amount of time to complete, even with only five participants completing three tasks, each five minutes long, so this would not be feasible for the main study with upwards of twenty participants. A method by which to code the observations was required. The primary and secondary researchers created their own informed categories of behaviour. Only behaviour relating to the agents or affecting the study were coded, actions to do with the task alone were ignored. *Figure 8* shows the comparison of each set categories created by the researchers. (A raw image of the whiteboard comparison can be seen in A*ppendix 5*)

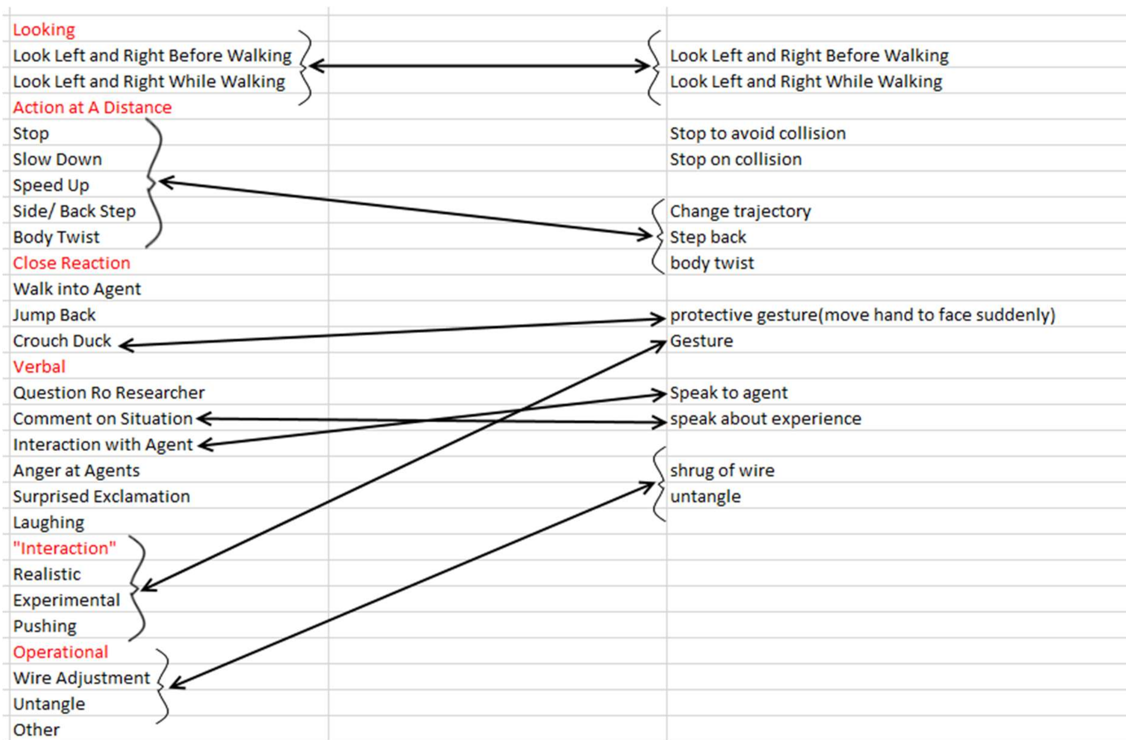


Figure 8: Summary of initial categorisation of behaviour by primary (left) and secondary (right) researchers. Arrows linking individual codes. Raw image in Appendix 5

Figure 8 shows clear differences in coding methods, which was to be expected due them being made independently. Discussion between the two researchers and another third researcher was then done. Topics discussed included whether to merge the looking codes into one or if the walking into agents can be accurately coded due to not knowing the participant's intent. Eventually, a final set of codes were devised and given specific descriptions. Six categories of codes were required, Looking, Action, Reaction, Gesture, Physical Interaction, Verbal and Operational. Each of these had multiple specific codes for certain behaviours, for example, slowing down for agents or experimentally touching an agent.

Once this was completed both the primary and secondary researchers coded a random participant from the main study independently. This was to ascertain whether they each interpreted the codes in the same way. Both researchers then compared these transcripts together whilst watching the participant and the differences were discussed. The differences at this point were substantial, with many codes being used for different reasons by each researcher. The main cause for this was edge cases where the behaviour seemed to fit multiple codes or none. Here is some example of problems that arose and the solutions to them:

- A Participant flinching did not have a specific code due to the behaviour being missed during the pilot, as a result, one researcher coded it as a "protective gesture" and the other as "reaction stop". To alleviate this a new code was added "reaction minor" to cover the behaviour.
- Similarly, the participant selected had a specific gesture not seen in the pilot. They would move their hands away from agents whilst keeping their body the same. One researcher coded this a body twist whilst the other frequently missed it. After some deliberation, a new code was added "avoidance gesture" to cover this scenario.
- "Wait" and "stop" were frequently coded differently, with one researcher using "stop" when the participant stopped to let an agent walk past and the other using both "stop" and "wait" for the same scenario. After some deliberation, it was agreed that they could be used together if a significant amount of time (e.g. 2+ seconds) passed from when the participant stops to when they carry on, there also had to be an agent or group clearly visible for them to be waiting for.
- "Watching" had a similar issue, with one researcher recording a "watch" while the other recorded a "look left/right". This was solved by requiring a significant amount of time (e.g. 2+ seconds) and a specific agent or group of agents for the participant to be watching.
- Another disagreement was when the participant slows down to a stop. One researcher recorded as "slow down" while the other recorded "slow down" and "stop". It was agreed that if the participant slowed and then stopped suddenly both codes were needed but if the participant slowly walked to a halt then only "slow down" was required.

This process was repeated a second time with a different randomly selected participant. In comparison, fewer differences were found with all major events being coded near exact. At this point, it was agreed that more comparisons would result in diminishing returns and so the final codes were agreed upon. A complete list of these codes with descriptions can be found in *Appendix 6.*

It was also agreed that if new, interesting behaviours were observed which could not be adequately coded, both researcher would discussion it possibly add a new code. No new codes other than "reaction minor" and "avoidance gesture" were added, however.

Medium	Code
00:04	START
00:24	Ast
00:24	Lw
00:24	Aw
00:34	Acd
00:53	Llw
01:04	Ast
01:07	Lw
01:07	Aw
01:11	Llw
01:18	Asd
..	..

Table 2: Excerpt from final codes. Note that multiple codes were possible at the same time.

This discussion between the primary and secondary research with input from a third was to minimise any possible biases. If the primary researcher coded everything with no discussion they could have missed categorising or failed to code specific behaviours and invalidated the work.

Once completed the codes for each video were counted and then merged with other participant's codes from the same density. The results and discussion from this data analysis can be seen in *Section 5.6* and *Section 6* respectively.

3.2.3 Qualitative Interviews

The final measure used was a series of semi-structured interviews, these provided some qualitative data to help support the quantitative data above. By structuring the interviews to promote free discussion, common trends started to become apparent. The analysis done with the interview data was limited, a full thematic analysis was not completed. This was because the interviews were not the primary source of data and were only done to support existing data with quotes.

The interviews consisted of six questions along with an open-ended question at the end. The aim of the project is to see how people react to a VR crowd simulation, so the first question was:

"How would you describe your feelings throughout the experience?"

This along with prompting follow up questions relating to the different stages aimed to support the PANAS data by allowing the participant to describe what the main influences were on their emotions. To follow this, they were asked:

"How do you feel about crowds in real life?"

This aimed to get the participant thinking about the simulated crowds' realism before being explicitly asked. It also allowed for any clear differences between how the participant think they react to real crowds and how they reacted to the simulated crowd to be compared giving insight into differences between them. This was further supported by:

"To what extent do you think that the virtual peoples' behaviour was realistic – was it like being in a real crowd?"

By asking this, it allowed them to compare the virtual crowd to real crowds. The aim was to discover any major aspects that changed the participants' affective state. Probing questions such as "was there anything that stood out as being strange?", or "do you think this changed how you interacted with the crowd?". Worth noting is that the word "agent" was avoided when talking about the agents, this due to its technical nature, that may have different meaning depending on the participant's background, "people" or "virtual people" were used instead to help ensure there was no confusion. A follow-up question which was sometimes unneeded depending on the previous questions answer was:

"Do you feel that the people were aware of you, and responded to you?"

Aiming to find out whether the participants thought they were being avoided, as the simulation model dictates or being walked over. This allowed the social forces model to be scrutinised as if they felt there were not aware of them like in a real crowd then further work in the future would need to be done to improve the model. Closely linked, was the next question

"Did you try to interact with the people?"

It provided insight into how realistic the participant's thought the agent's behaviour was on an individual level. The final structured question was

"Apart from the restrictions of the Vive, did you find it difficult to move around the environment (more difficult than in a real crowd)?"

This aimed to discover whether the participants found it harder to move around in the virtual crowd. Prompting questions relating to whether it changed between conditions were also asked, to see if density had any impact. The Vive restrictions were avoided as it was not the focus of the study, steps were taken to minimise its effect explained in *Section 5.1.2* and *5.2*. The final open-ended question:

> *"Is there anything you'd like to say or ask which you haven't had a chance to mention?"*

Was used as a catch-all to allow the participant chance to say anything they wanted about the experience and the study. It also allowed them to ask questions that they may have.

4 Design and Development

This section aims to explain how the crowd simulation was implemented and validated, and then how it was adapted to allow a user to interact with it in a virtual space.

4.1 Unreal Engine

As discussed in *Section 3.1*, Unreal Engine (Epic Games, 2015) was used to create the application. The main unique feature used within Unreal was the Behaviour Tree, a finite state machine used to control the agent's actions and variables. It allows intuitive control over the agent's behaviour. Variables such as the agents current objective and despawn points are managed here.

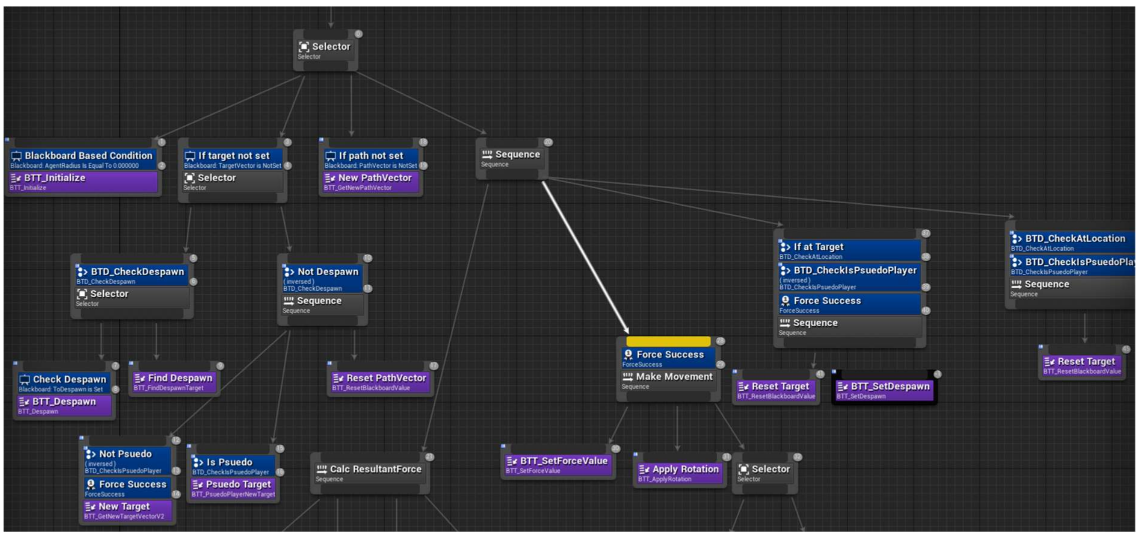


Figure 9: Unreal's Behaviour Tree system.

Another feature used was the Blueprints, a visual scripting language. These were used to make the non-computationally intensive sections of the application, such as setting the current desired destination or applying rotation to the agent. These Blueprints are not able to run as effectively as C++, meaning expensive procedures such as each agent finding the relative distance of every other agent were programmed in C++. Forces were also handled in C++ as they had to be applied to the agents every frame which the Blueprints and Behaviour Tree could not do.

4.2 Crowd Simulation

4.2.1 Unreal's Detour Crowd

The initial implementation for the crowd simulation used Unreal's (Epic Games, 2015) crowd management system, Detour Crowds. This, like other crowd simulation models, is designed to allow the agents to navigate around a map while avoiding obstacles and other agents. This quick implementation allowed the researcher to design the other features of the application such as how each agent decides its current goal and when and where they spawn and despawn. By doing this before implementing the full social forces model (Helbing and Molnar, 1995) it allowed for easier implementation, as the expected agent behaviour is known. Once the agent management and individual decision making were implemented the social forces model was implemented, replacing detour crowds as the method of navigation.

4.2.2 Helbing's Social Forces

As presented in *Section 2.1.1*, Helbing's social forces model (Helbing and Molnar, 1995) is made up of three force categories, the driving force, avoidance forces and grouping forces. This force based model makes the conventional, velocity based AI used in Unreal unfeasible. This meant that any of Unreal's existing navigation methods would override the simulation model, nullifying its existence. This means some of the basic features Unreal offers such as its navigation and behaviour handling could not be used, requiring a bespoke navigation system to be implemented. For this reason, the agents were created to be physics objects, able to be pushed around but not be influenced by Unreal AI systems directly. During development, basic cylinder models were used to represent the agents, they had weight but were locked in the Y direction a millimetre above the ground so did not have any surface friction. This constraint meant that the implemented model could not be used to navigate across the Y direction, such as stairs or slopes, this was not required for the study, however.

4.2.2.1 Navigation (Driving Force)

The first force implemented was the driving force, used to push the agent around static objects such as walls. By using Unreal's pathfinding, path vectors can be extracted which the agent can follow concurrently to reach its objective (*Figure 10A*). By calculating the unit vector between the agent's current location and the next path vector, the current desired direction of travel can be established. Then, since the Driving Force is constant the unit vector can then be multiplied by a fixed magnitude and applied as a force to the agent. By varying the size of the constant and the agents linear damping the maximum speed and acceleration can be controlled.

This agent is now able to navigate the map and reach its desired objective. The problem with this, however, is the agent must reach the exact position of the path vector which is unlikely to happen with other agents possibly pushing it around. For this reason, an acceptable radius around the vector was added allowing the agent to be within a 0.2m margin of error to succeed. The produced another problem however with the agent sometimes apparently reaching its goal when the next path vector is not in direct line of sight (*Figure 10B*). To fix this, instead of the agent continuing to follow the previously made path, it will create a new path ensuring it will not try to walk through obstacles.

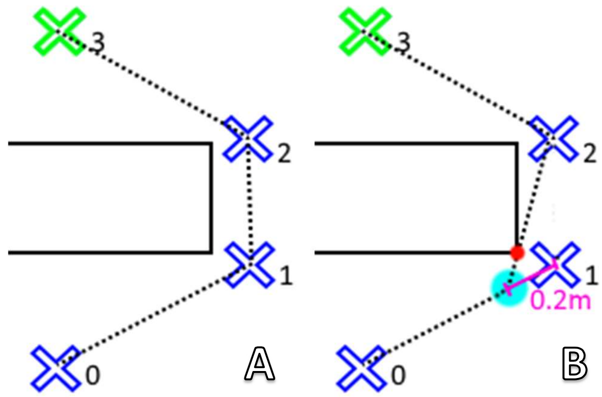


Figure 10: A: Path vectors showing current, 0, and target, 3, location along with other path vectors required to avoid the static objects. B: possible collision if a new path is not created when the agent is within range of the vector but not past the obstacle.

4.2.2.2 Inter-Agent Forces (Avoidance force)

The agent is now able to move from point A to point B. The agent still has no sense of other agents and the pathing around obstacles looks very artificial with the sudden turns. This is where Helbing's other forces are used.

i $$f_{ij}^{interact} = A_{ij} * \exp\left[\frac{R_{ij}-d_{ij}}{B_i}\right] * n_{ji}$$

The equation above shows the first avoidance force implemented, the interaction force, which describes how 2 agents *i* and *j* interact over a distance, d_{ij}, with their combined radius, R_{ij}. The constants A_i and B_i denote the strength and range of the force respectively. This then points along the normalised vector between *i* and *j*, n_{ji}. As the distance between the two agents decreases the force increases pushing *i* away from *j*.

ii $$f_{ij}^{body} = k(R_{ij} - d_{ij}) * n_{ji}$$

iii $$f_{ij}^{slide} = c(R_{ij} - d_{ij}) * \Delta v_{ji}^{t} t_{ij}$$

If the distance between the two agents, d_{ij}, becomes equal to or less than their combined radius, R_{ij}, they are touching. If this happens, two new forces are applied. ii denotes the body force which applies the large force of magnitude k. This keeps agents from passing through each other or simulate pushing in a panicked crowd depending on how big k is. Along with this is the sliding force, *iii*, which pushes along the tangent to direction of motion, t_{ij}, and tangential velocity difference,Δv_{ji}^{t}. This force helps push two agents past each other as if they were to make a side step around them.

iv $$1 + cos\left(\frac{\theta * \pi}{180}\right)$$

The interaction force, *i,* is also multiplied by the function, *iv*, which makes agents in front of them influence their behaviour more than those from behind, with θ being the angle (degrees) from directly forward.

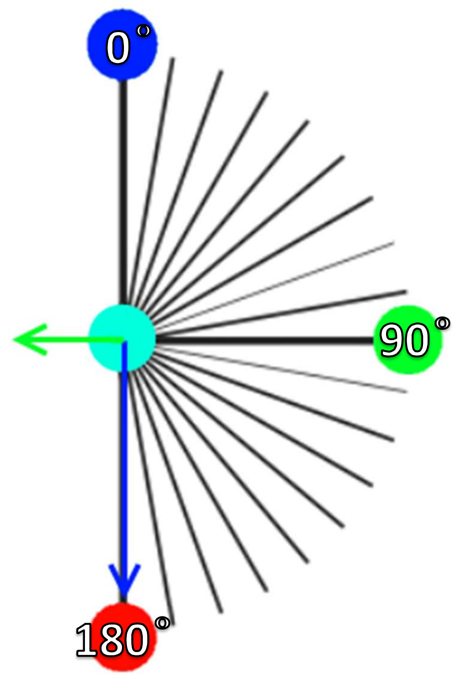


Figure 11: Showing an agent facing up the page with 3 other agents at 90˚ increments and their relative resultant force on the agent depending on their position.

The sum of these 3 forces completes the inter-agent forces within Helbing for this project. There is another force that could have been implemented, the grouping force. This allows groups of agents to be pulled towards each other simulating friends and family wanting to walk as a group. It was an unnecessary complexity, however, and not required to complete this project's aim.

The initial experiments indicated that under some cases the standard model's avoidance force was not adequate to prevent obvious collisions. This is particularly the case in near-head-on collisions, where the force acts near-parallel to the direction of motion, and there may be too little perpendicular force to prevent an overlap (especially in this case, where fully animated body representations are used for the agent). A similar phenomenon has been reported by Steffen and Seyfried (2008). In some previous works, collision prediction has been used: in this case, an additional force (which is referred to as "Turning Force") is introduced to counter the specific case of head-on approaches.

The magnitude of this force is relatively small, and scales inversely with distance, such that agents moving directly towards each other turn slightly to avoid head-on collision whilst still separated by several 10s of metres. The force is negligible over a range of several metres, as smaller perturbations are required as the magnitude of the repulsive force increases.

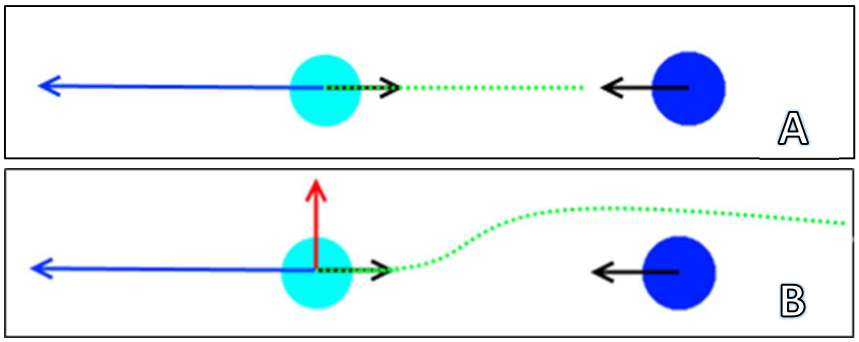


Figure 12: Image showing an imminent collision between 2 agents with (A) and without (B) a turning force. Green = trajectory, blue = avoidance force, red = turning force and black = current direction of motion.

v $$f_{ij}^{turn} = g * (R_{ij} - d_{ij}) * \frac{\theta}{\sqrt{\theta^2}} * p_{ij}$$

Here the "Turning Force" is described. Parameters are much the same as previous equations with 2 main additions, p_{ij} is the normal vector between *i* and *j* rotated 90 degrees around the y-axis. Assuming θ isn't 0, $\frac{\theta}{\sqrt{\theta^2}}$ is the angles between them reduced to 1, this dictates the direction of force be it left or right, if it is 0 it is assumed to be 1. This force is active when the other agent is within max avoidance range (100m, *Table 3*) and inside ±15˚ of the forward vector (centre 30˚ of field of view). Along with this the sum of the two agents' forward vectors must have a magnitude less than 0.5, this ensures agents don't try to turn away from agents walking away from them.

4.2.2.3 Obstacle avoidance forces

The object avoidance forces are analogous to the inter-agent forces, with the distance being between the agent and the closest point on the obstacle, the turning force pushes tangent to the object in the direction of travel. The inclusion of these forces not only allows agents to avoid objects not pathed around but also makes cornering around walls smoother with less sharp turns, as seen in *Figure 7*, improving realism.

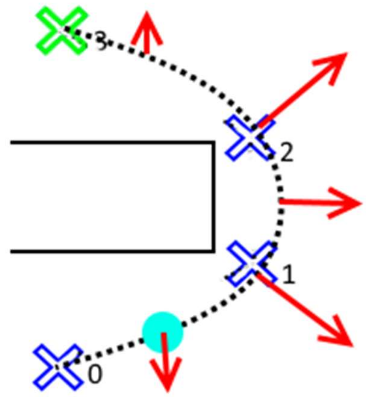


Figure 13: Obstacle force (red) while walking around a wall. Notice how the path is curved and does not necessarily pass through the path vector origins.

4.2.2.4 Error Handling

At this point, many agents can move through the map with their own individual objectives avoiding one another and objects around them. Problems around the path vectors did arise, however, if an agent walks past a path node without passing within the threshold distance (0.2m) for example, they would turn around and walk back towards it, even if they would be walking away from their final objective (*Figure 14*). This sometimes-caused dense crowding on corners unrepresentative of the whole simulation. To counter this the distance from the next point is recorded and updated every frame so if it ever increases, suggesting agent has been pushed away, a new path is created.

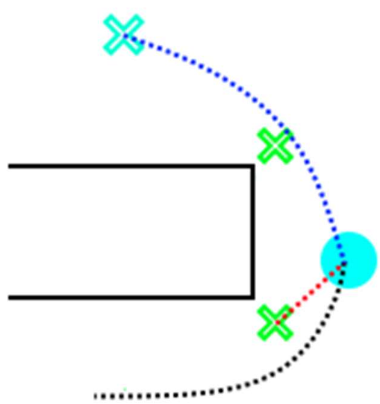


Figure 14: Agent pushed past path point without hitting it, with uncorrected path in red and desired path in blue.

Another issue arose in that path vectors around walls would always form in the same spot (*Figure 15A*), on the corner of the navigation mesh, to help counter this another spot on the mesh was chosen at random (within 0.25m) of the ideal vector (*Figure 15B*). Although not perfect it did add some variance to agents cornering around objects.

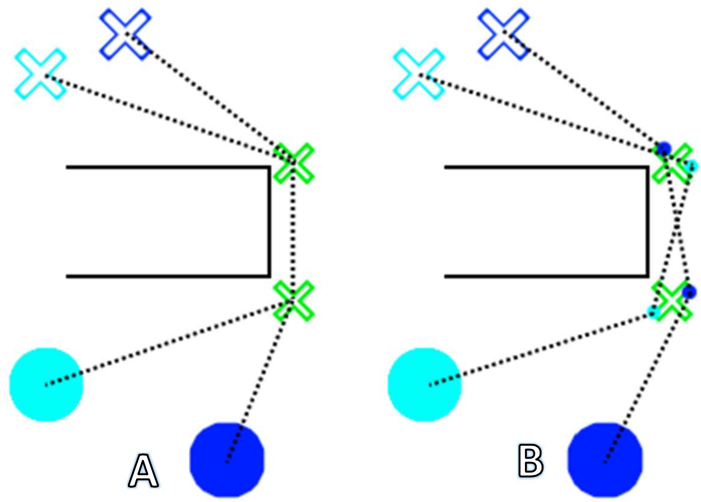


Figure 15: Path vectors showing desired path with (B) and without (A) randomisation.

4.2.2.5 Force Values

Attribute	Value
i	
A (Avoidance force)	100
B (Avoidance Max Distance)	100m
ii	
k (Body force)	150
iii	
c (Sliding force)	1600
v	
g (Turning force)	3.33
Force Weighting	
Driving Force	0.4
Agent Avoidance Force	0.6
Obstacle Avoidance Force	0.7
Agent Properties	
Linear Damping	14
Weight	1kg
Agent Radius	20cm

Table 3: The force constants and weighting along with the agents' relevant physical properties.

Table 3 presents the values used for each constant. These were determined through trial and error by using the developer's subjective opinion, these are validated in the next section. In the above table is the weightings used for each type of force are also present, these could for agents' to have individual personalities as discussed in *Section 2.1.1*, in this case however, all agents had the same weighting and were solely used to further calibrate each force. The relevant agent properties are present also, linear damping being the sole form of friction as the agents did not make physical contact with the floor.

4.2.3 Validation

Realistic crowd movement within a simulation is defined from the emergence of group behaviour consistent with those observed in real-life scenarios (Pelechano et al, 2007a, 100). This causes some problems when trying to validate the model, as there is currently no way to measure group behaviour in virtual crowds. For this reason, a lot of crowd simulation models rely on intuition and assumptions on the part of the developer to calibrate parameters, rather than scientifically validated observations.

An objective method by Schadschneider and Seyfried (2009) used a fundamental diagram (*Figure 16*; Seyfried et al, 2005) to verify their simulation. The diagram plots the measured speed of real people walking around a narrow lane, against their crowd density. This allowed for an objective comparison to be made between real people and their simulated counterparts.

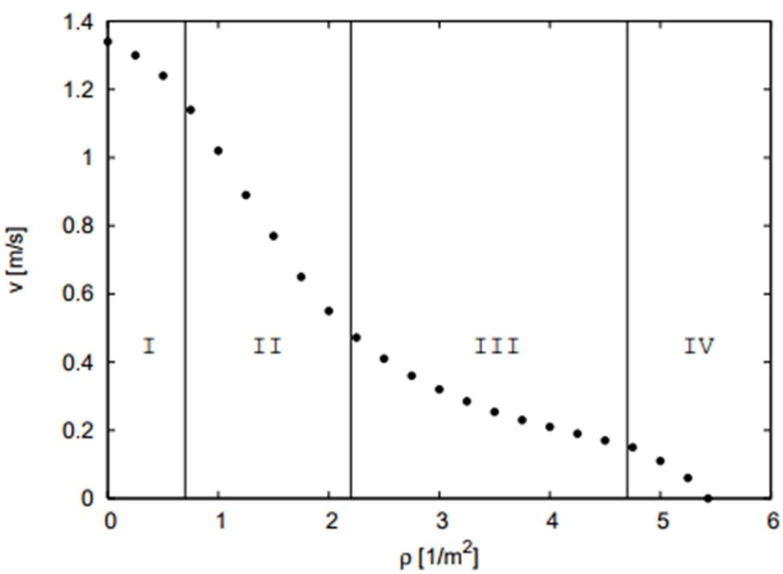


Figure 16: graph showing the relationship between Velocity and Density of real people. (Seyfried et al, 2005)

A similar method of validation was used for this project. As shown in *Figure 17,* a test scene incorporating a large circular corridor was created; allowing hundreds of agents to traverse around it. A 2mx2m section (shown in red) was placed in the middle of the corridor to track the average velocity and density every 0.5 seconds that an agent was walking through it.

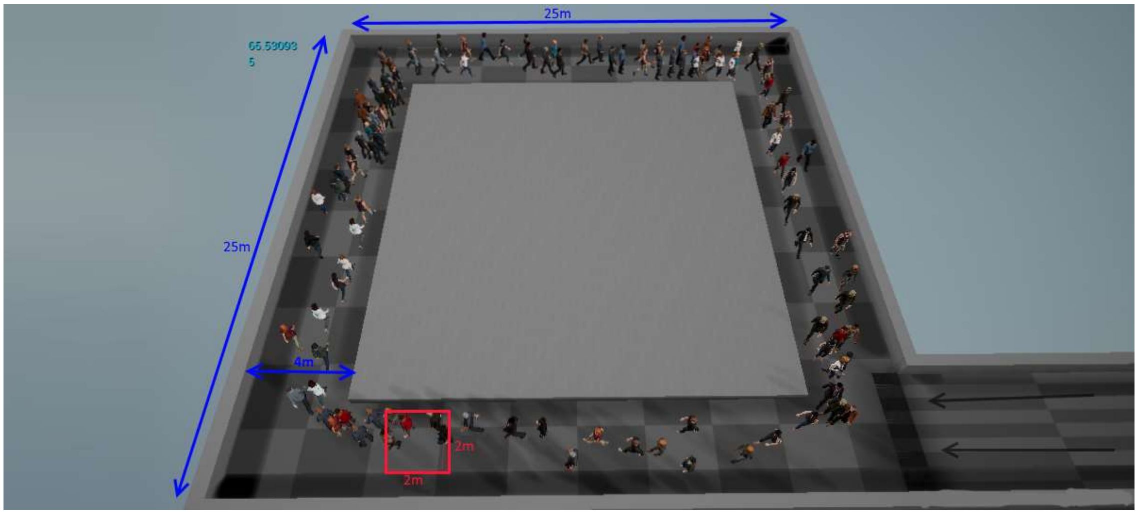


Figure 17: Corridor map used to create the Velocity v Density fundamental diagram seen in Figure 18.

This was run for a total of 30 minutes across 3 instances for a total of 1276 data points. An important step taken whilst running the simulation was to allow the agents to walk onto the map naturally (the entrance bottom right on *Figure 17*) as it is difficult to initialize the scenario with realistic starting positions for all the agents, this meant by the time they reach the tracking area they had spread out and walk as the model dictates. The output of this log (*Figure 18*), when compared to the same graph with real people (Seyfried et al, 2008) show the form and range is comparable.

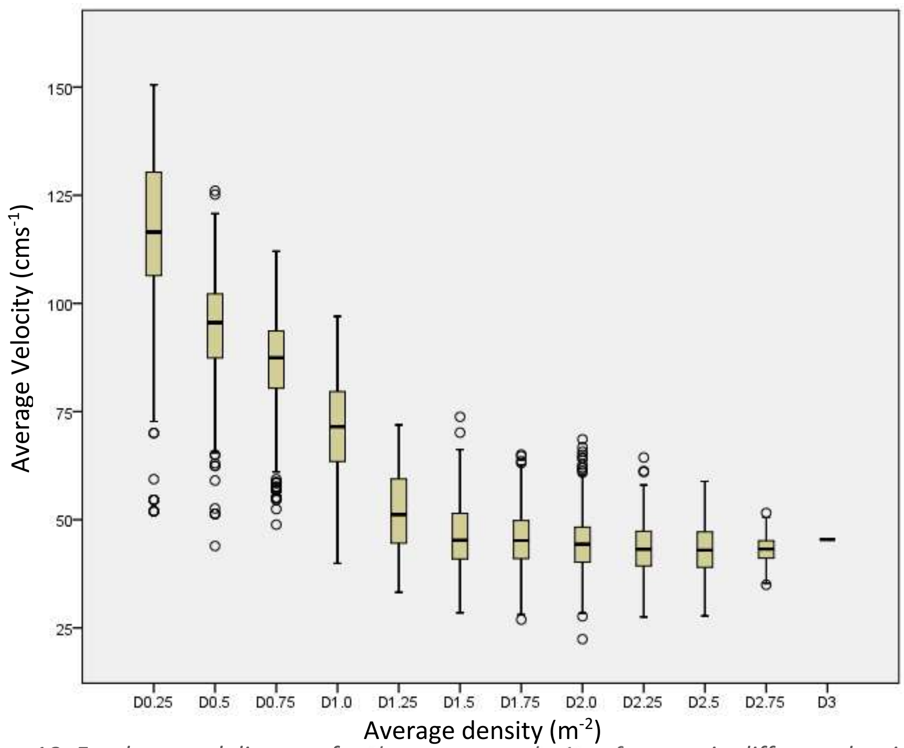


Figure 18: Fundamental diagram for the average velocity of agents in different densities. (N = 1276, t = 1595s)

By comparing the fundamental diagram with the one created using real people (*Figure 16*) the general shape and values can be seen to be comparable, this along with the general realism in the observed behaviour presents good evidence for the simulation to be validated.

4.3 User Interaction

The section will outline the steps taken to allow the user to interact with the simulation. This includes character models and additions made to the simulation allowing the agents to react to the user in a realistic way whilst still following Helbing's social forces model (Helbing and Molnar, 1995).

4.3.1 Character models

At this point the agents have been represented by cylinders with arrows on top representing direction of travel which was perfectly acceptable for a top-down view used to build and validate the model but when interacting with it in a first-person perspective the ability see realistic models and animations increased immersion, it can also help in further validating the simulation (Pelechano, et al, 2008b). The creation of these models and animations would take a long time to make by hand, so Fuse's (Adobe Systems, 2014) default assets were used to create the initial 10 models (5 male, 5 female). The animations used for each model were also varied by using the multititude of available animations created by other Fuse users. The appearance of these assets were selected for their suitability in a university environment.

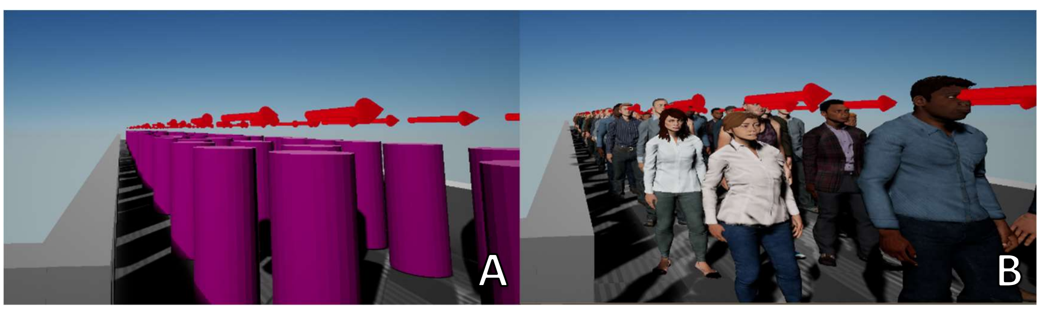


Figure 19: Change in Agent appearance from cylinders to character models. Trajectory arrows were kept until late in development to help with debugging.

4.3.2 Agent-User Forces

The initial player character was Unreal's first-person character asset; this allowed the developer to see the world from a first-person perspective without having to use VR to do so. Allowing the agents to perceive the user was a relatively simple task. By attaching an invisible agent to the player character all other agents could react to the character in the same way they do themselves. The initial aim was to have the agents interact with the user in the same way as they would any other agent. This quickly proved ineffective however as the agent would not react enough to avoid the player. This was because when two agents are avoiding each other they experience symmetric avoidance forces. With a human-agent interaction, only the agent is receiving a force so only half the resultant forces are applied. It could be argued that this is the correct method as the user will be trying to avoid them also but in practice, it did not work, for example when the user is stood still the agents often collided with them. To counter this the avoidance forces from the user on the agent increased by a factor (in this case, the value of two, estimated by trial and error) to help simulate the agent-agent interactions. With this, the user was able to walk around the game world freely with agents avoiding them. The character models were also checked at this point to make sure they were of realistic height compared to the character and world around them.

It became clear at this point that if the user tried to take advantage of the simulation and knew how it worked they could easily cause unrealistic behaviour in the agents, such as getting them to walk into walls and obstacles by pushing them and trapping them in a corner. With real people, they could push past the user but the in-game agents have no way of imposing a force upon the player. If the user walked around the world as they would in real life however, the system worked as intended.

4.3.3 Virtual Reality

To implement VR, Unreal's motion controller asset was used, this allowed the user to interact with the world within a 2m x 2m area. To counter the small area the user could teleport to new locations. Early in development, the aim was to make the simulation as realistic as possible however, this meant the ability to teleport was disabled. Because of this, the user will be restricted to within an area the size of the real world Vive area. The limitations of this will be talked about more in the Scenario section of the evaluation (*Section 5.1*).

5 Evaluation

With the VR crowd simulation fully implemented the study could commence. This section will discuss the decisions made on how the study was run and explain what the participants did at each stage. It will also explain any aspects discovered during the pilot that could be improved and changes made to the system and procedure before the final study began. The results will be analysed and presented to then draw some conclusion as to how the participants reacted to the system.

5.1 Scenario

5.1.1 Environment

When creating the environment, the aim was to ascertain how users responded to the simulation. To do this an experimental scenario based on a real campus building at the University of Lincoln was reconstructed from exact blueprints of the building. The area is a busy thoroughfare, and typically has a lot of bi-directional pedestrian traffic. It is also used by charity stalls to sell cupcakes, and by schools to advertise courses available to prospective students along with other people interacting with any passers-by. In this scene, virtual characters are spawned at set increments at either end of the thoroughfare and move to a point at the opposite end. Start and end points are beyond the view of the participant, such that this appears as a natural flow of pedestrians. This meant the map was a lot larger than the small area that the user can use as permitted by the Vive. The ability to see the whole area and see agents as they approach aimed to give user's time to observe the agents before having to react.

5.1.2 Lab setup

The key requirements for the study set up were for the area to be big enough for the user to complete a task that required them to walk parallel and perpendicular to the flow of virtual agents. A method of lowering the chance that the participant will become impeded by the wire and a way to easily see and record what they are doing while in the simulation and link it to any gestures or verbal queues from them.

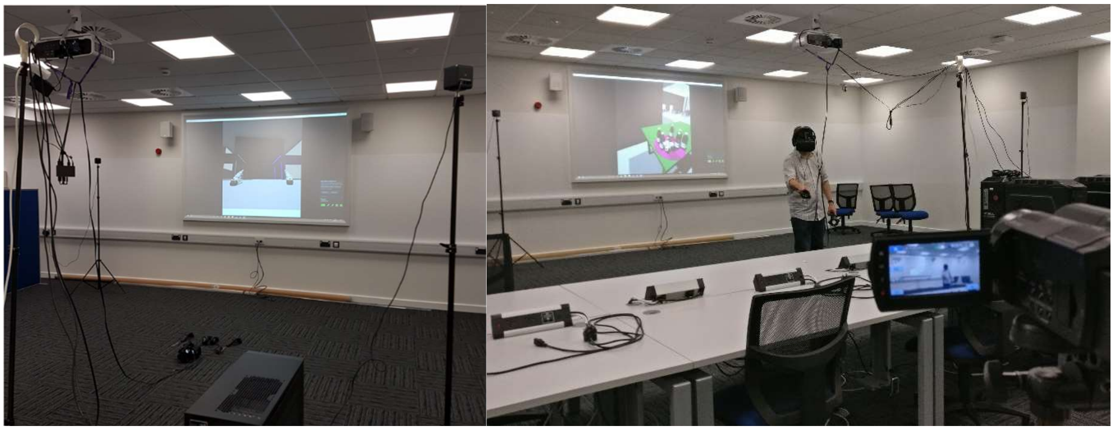


Figure 20: The lab set showing projector and camera along with wire suspended above the user to avoid tangling

To achieve this, the study was held in an unused computer lab with a large floor space and projector, as seen in *Figure 20*. This allowed for the play area to be in front of the projector screen so anything the participant sees can be easily seen by the researchers and recorded. The projector, being near the centre of the area, also provided a place to anchor the cable above the participant's head while still allowing the cable to reach the entire area. To avoid any intrusions during the study a sign was placed on the door instructing people to wait outside until the study had concluded.

5.1.3 Task

Participants were asked to play the role of someone setting up some tables in the thoroughfare. They were tasked with moving 20 items, 10 bottles and 10 cans, to the green and blue tables respectively. To do this they need to walk perpendicular to the flow of agents and parallel in both directions. Participants thus interact with agents while performing a repetitive manual task.

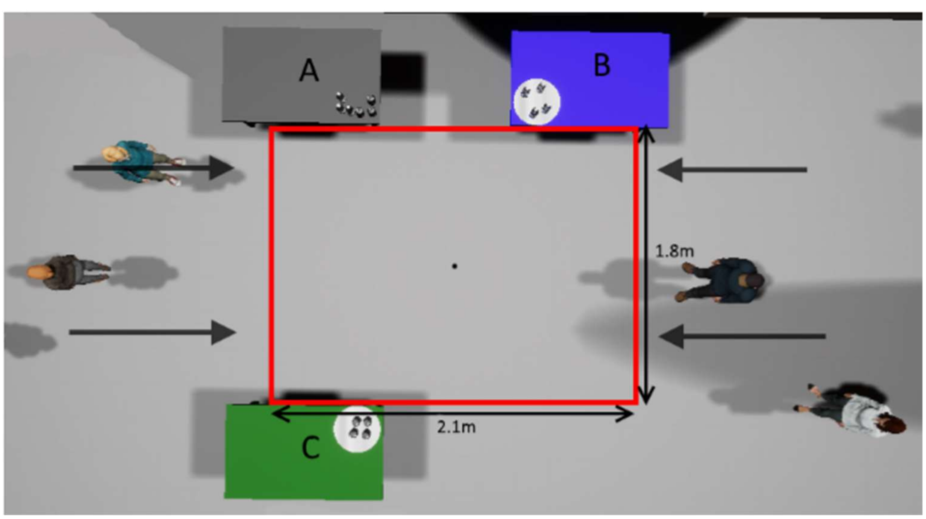


Figure 21: Participant play area showing the items on the 3 stalls along with the directions the agents will be flowing. Note white circles used to detect when task is complete.

The task can be completed in any order; however, when they were trained to do the task, before undertaking the experimental conditions, they are told about the bottles to the green table first. This meant most participants chose to move a bottle first even when told they could complete the task in any order. To keep the task similar for people who were used to VR and those who were not, only one object could be picked up at a time.

5.2 Procedure

The Participant was first welcomed into the room allowing them to clearly see all the equipment, they were then told a summary of what was going to happen. They were also introduced to the second researcher who would be operating the camera. Once they agreed to proceed they first filled out a medical screening form to make sure they could safely take part and not have any related problems, such as uncorrected vision, fear of crowds or claustrophobia etc. At this point, they were assigned a unique ID, this would be used to identify their information anonymously by referencing it in on any data from them along with the consent form which is kept in a separate secure location and is the only place that references both their name and ID together; this would not be used during analysis. They were asked to sign this consent form which explained everything that would happen and there right to withdraw at any point. A demographics form was completed along with the first PANAS form, it was explained to them that it is was a validated questionnaire used in many studies in the past and that they would be filling out 5 forms in total. This introduction took between 10 – 15 minutes to complete, discrepancies in time were due to some participants asking more questions and needing an explanation with some of the forms.

The participant was then led over to the play area where the HTC Vive was presented to them, including an explanation of what it does and how it works. The chaperone system, warning players of the edge of the area was explained to them before putting the headset also. They were reminded that a camera would be recording them whilst in the Vive; this was explained to them earlier in the participation sheet, and during the introduction. Once the headset was on and adjusted correctly the participant was in the default training environment for the Vive when no applications are running. Whilst in the environment they were asked to walk around the outside of the area using the chaperone to assist them. This was to make sure they were confident with the size of the area and that there was nothing to bump into. They were also warned about the wire and how it may impede their movement with advice about how to move it if required. While in the simulation, the researcher lifted the wire over the participant’s head whenever necessary, the participants were not informed of this beforehand as It may have affected their behaviour during the experimental conditions. During the interviews they were asked if they were aware of the Researchers presence at any point during the study, none of them were.

Once the participant had a few minutes to become acclimatised to the situation and were comfortable moving around the empty map they were taken to an empty scenario map with no agents, the camera was turned on. This map was used to talk them through the task and explaining that it was not a race and they will not be judged by how well they do. They were also told that if an item falls outside of the playable area that they should not try to walk beyond the chaperone system and to get it and to just carry on moving the rest of the items. This training section took between 5 and 10 minutes to complete depending on how easily the participant interacted with the world. Adjusting the Vive to fit each participant sometimes took longer than expected. Once they had moved all the items to the correct places they were asked to take off the Vive headset and fill out another PANAS form.

With this completed they were then asked to return to the Vive and told that the researchers would not interact with them this time, to help with this the participant would wear a headset, they were reminded however that the researchers will still be able to hear them so if they wished to stop or needed help they could call out. They were also reminded of what the task was once again.

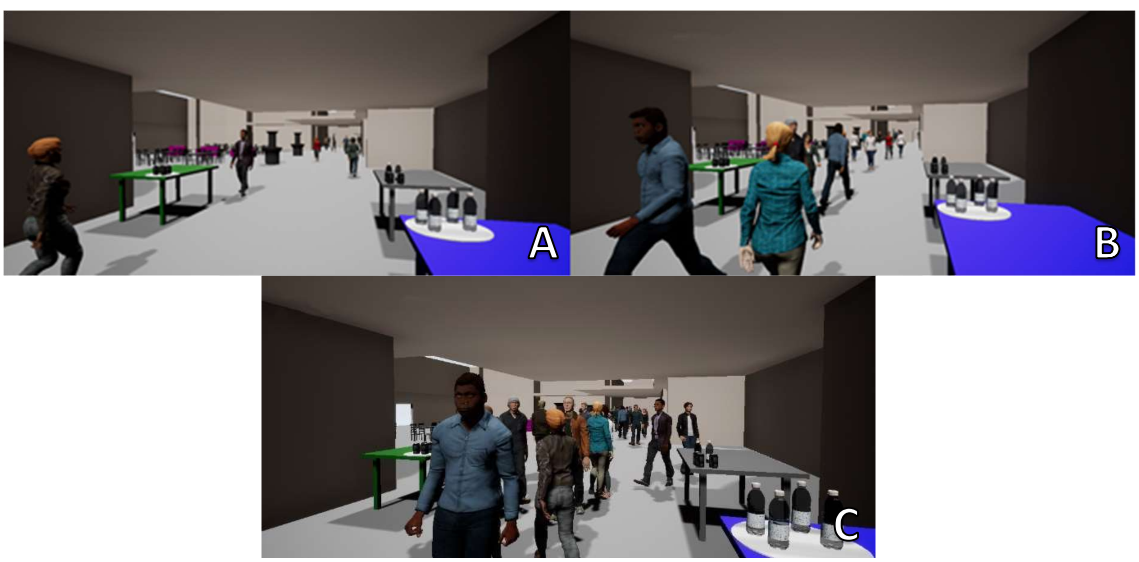


Figure 22: Screenshots of the scenario showing the three densities Low, Medium and High that the participant interacted with; labelled A, B and C respectively.

The experiment comprised 3 conditions (*Figure 22)*: the participant completed the same task in each, and the crowd density was varied as an independent variable. The average densities for each condition can be seen in *Section 5.5, Table 4*. In practical terms, this meant low density allowed the participants to walk across with little chance of collision, in medium the participants would need to make some adjustments in the movement to avoid them but have little difficulty and high being a constant flow of agents in which the participant would most likely have to push past some agents to cross. These conditions were counterbalanced between participants using a Latin square to avoid any ordering effects. In total, each condition took 2-3 minutes to complete. Participants were then asked to complete a PANAS questionnaire at the end of each condition.

Once the final condition was completed a 5- to 10-minute semi-structured interview was conducted. The interview was recorded using a voice-recording device and subsequently transcribed by the researcher. This provided additional qualitative data about the participants' experiences while completing the conditions and to allow them to input and ask about any thoughts during the study.

In total the study took between 30 to 45 minutes to complete, under the time specified in the participation sheet to capture any outliers taking more time. The section with the largest fluctuation in time were the interviews due to the differing response from participants.

A copy of the medical and demographics forms along with a PANAS questionnaire, consent form with information sheet can be found in *Appendix 1-4.*

5.3 Pilot Study

Before commencing with the main study, a pilot was held to check if the current procedure and scenario were going to be able to fulfil the stated aims. In total six participants took part, with a 50/50 split in gender with an average age of 24.3 (standard deviation: 9.8). Four of these participants were students with two being staff. Half of the participants had used VR in the past.

5.3.1 Revisions

After the pilot was completed it became clear that a few changes had to be made to the simulation and procedure. Below these changes are presented with reasons for doing so.

5.3.1.1 Trajectory Log

An increased timer resolution would provide better quality data, by only logging a tenth of a second at twenty logs per second every two logs had the same time, to fix this a milli-second timer was implemented. Individual logs were also added for each hand allowing accurate tracking of when an object is picked up or dropped.

5.3.1.2 Character Models

During the pilot interviews, two participants commented on seeing duplicate agents or "twins". This was because of the low number of models, to fix this the number of unique models was doubled to 10 male and 10 female, during the study the Researcher also noticed one of the agents model appearing more frequently than others this was because it was duplicated in the randomiser so it had twice the chance of being chosen compared to any other model. Care was taken during the creation of these models to make sure each one didn't look out of place in a university environment and were unique, by varying the facial features on the different models along with colour and material changes on their clothing.

5.3.1.3 System Automation

To run the study the Researcher had to remember which order the scenarios had to arrive in every time a scenario ended. To fix this the system had a level manager implemented which had the Researcher input the Participant ID along with the map order (*Figure 23*) at the start so the only thing the researcher had to press between scenarios was a single button. Because this was projected onto the projector screen clearly visible to the participant the names for different densities were numbered so the participant did not know what would be changing between maps.

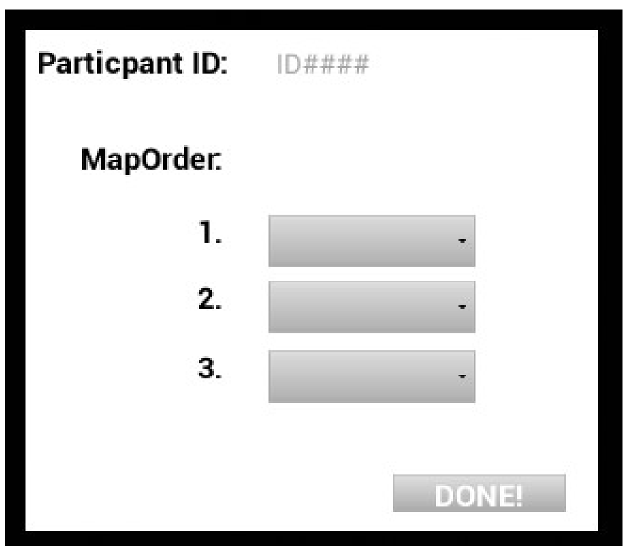


Figure 23: Screenshot of the initial data input screen used to save the participant ID and map order.

5.3.1.4 Other

Two participants mentioned how they found the task long and repetitive. in an attempt to mitigate this, in the final study, the number of objects was lowered to seven of each type, causing the minimum amount of times need to cross the area dropped from twenty to fourteen. During the study, some participants realised they could push agents with the items in their hands this was overlooked during the development phase and the ability to do so was taken out before the final study. This meant that no collisions (Physics based. People, objects and agents could still overlap) could happen between the player and agents or the items and agents.

5.4 Participants

Participants were gathered throughout the university campus by word of mouth. A Poster was designed and used in the last two weeks of the study as the flow of participants slowed. In total 25 participants took part with 13 male and 12 female, with an average age of 29.7 (standard deviation: 9.4). 18 of these were staff and 7 were students with 14 having used some form of VR in the past.

5.5 Virtual Participant

After the completion of the main study, it was decided that having an agent do the same task as the participants could provide useful insight into the differences between the behaviour of agents and human participants. This was implemented by programming a single agent walk between the tables as the people did while doing the task, i.e. fourteen times across and six parallel, whilst logging the trajectory in in the same way. The virtual participant was run a total of twenty-five times for each density, to match the amount of data gained from the main study, the trajectory data was then analysed with the same tool used to analyse the participant trajectories.

5.6 Results

This section describes the quantitative results from the collected affect, trajectory and video data is presented. To begin, the densities for each scenario were calculated empirically from the trajectory data from each participant. The agents within the walkable area were sampled 20 times each, for each condition and the counted (total of 500 samples per condition). The average density for each condition along with standard deviation are reported in *Table 4*.

Condition	Interval (s)	Density (m^{-2})	σ
Low	8	0.078	0.077
Medium	4	0.153	0.094
High	1.75	0.425	0.131

Table 4: Mean density of agents in walkable area across the three conditions. including the interval between new agents spawns

5.6.1 User Affect

Condition	Positive		Negative	
	μ	σ	μ	σ
Low	34.6	9.0	12.6	3.6
Medium	34.5	9.4	13.4	4.0
High	32.8	8.6	15.8	6.5

Table 5: Summary of PANAS statistics (Positive and Negative Dimensions). μ = Mean, σ = Standard Deviation

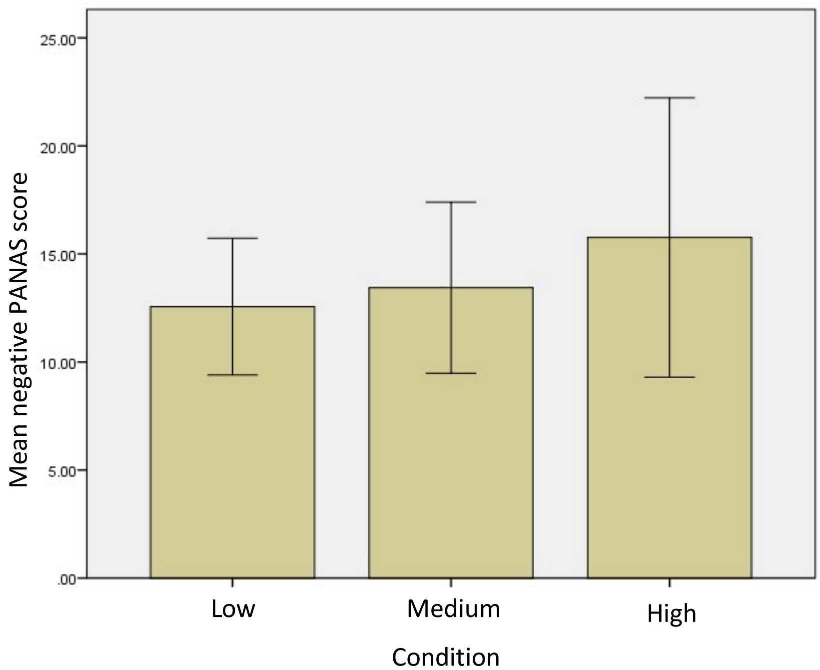


Table 6: Graph Showing mean negative PANAS scores, with error bars showing standard deviation.

The high-density condition reported higher mean values than the medium and low conditions, on the negative dimension and a lower mean value on the positive dimension. The data's sphericity had to be determined so Mauchly's test were conducted. The test for negative results indicated the assumption of sphericity had not been validated, $X^2(2) = 4.405$, p=0.111; the same was true for Positive, $X^2(2) = 5.071$, p=0.079. Upon consideration however, it was determined the small sample size negatively impacted the result, a common flaw with Mauchly's test (Laerde Statistics, 2013, 1). For this reason and the mean values showing clear differences between low-high vs medium-high, it was assumed that the data was not spherical. This led us to use the Friedman test to examine the statistical significance of these results, which showed significant effects on the negative dimension, $\chi2(2) = 13.975$, $p = 0.001$. Wilcoxon signed-rank tests were used as post-hoc, with Bonferroni correction setting the significance value at $p < 0.017$. statistically significant increases in negative affect were identified between the high and low condition (Z = -2.842, p = 0.004) and high and medium condition (Z = -2.725, p = 0.006), establishing an increase in negative affect from a higher crowd density. The Friedman test indicated an increase between low and medium but post-hoc analysis showed no significance. The null hypothesis H_0 (*Section 3.2.1*) is rejected accordingly.

The Friedman test indicated significant effect on positive affect $\chi^2(2) = 6.841$, $p = 0.033$. Post-hoc analysis, however, using Wilcoxon signed-rank tests (Bonferroni correction, significance level $p < 0.017$), showed no significant difference between low, medium and high. The raw PANAS results can be viewed in *Appendix 8*.

5.6.2 Trajectory Analysis

Trajectory data was used to further quantify the participant's activity and experience. To do this the data was post-processed to extract many metrics. On completion of the study, it was found that one of the participant's data files from one condition was missing; this participant was removed from the trajectory processing.

The first metric extracted was the extent to which the participants were impeded in their task by comparing their mean walk speeds whilst crossing the area. To do this the time spent stationary at each table had to be excluded. A central area was defined as 0.5 metres within the walkable area (in most cases this correlated to the participant walking across the area). The speed was then computed from each visit into the area. The mean velocities for each participant, for each condition, were calculated. A summary of this data is in *Table 7*. A repeated measures one-way ANOVA showed significant effects between conditions at $p < 0.05$ ($F(2,48) = 42.525$, $p < 0.0005$). Post-hoc test with Bonferroni correction showed significant differences between low- and high-density conditions ($p < 0.0005$) and medium and high conditions ($p < 0.0005$).

Condition	Speed (ms^{-1}) Mean (μ)	Speed (ms^{-1}) Standard Deviation (σ)
Low	0.74	0.10
Medium	0.71	0.12
High	0.61	0.12

Table 7: Mean and standard deviation for walking speed while crossing flow of virtual agents.

Previous research, discussed in *Section 2*, looked at the proximity of agents to participants; so, to more closely link this project the closest average proximity the other agents was calculated, this is summarised in *Table 8*. For this, samples were taken at regular intervals for each participant in each condition (120 per condition). For each sample, the distance to the nearest agent was computed. The mean distance was then computable along with standard deviation. A repeated measures ANOVA showed significant effects between conditions ($F(2,46) = 41.339$, $p < 0.0005$). Post-hoc test with Bonferroni correction showed significant differences between all conditions at $p < 0.05$. *Section 6* discusses how this connects to previous works in VR proxemics.

Condition	Distance to nearest agent (m)	
	Mean (μ)	Standard Deviation (σ)
Low	2.08	0.09
Medium	1.57	0.13
High	1.05	0.09

Table 8: Mean and standard deviation for agent proximity to participant

5.6.3 Observational Data Analysis

In total 28 codes were created which fit into broader categories, as shown in *Table 9*. A complete list with descriptions can be found in *Appendix 6.* Three participants did not get recorded properly with 1 or more of their videos becoming corrupted or lost due to miss management, these participants were excluded in this part of the analysis. 66 videos (22 participants) were still valid.

Category	Codes
Observational	Look While Stationary, While Walking, Watching
Actions over a distance	Sudden Stop, Slow Down, Speed Up, Change Direction, Wait
Reactive	Stop, Collision, Step Away, Protective Gesture, Avoidance Gesture, Body Twist, Minor Reaction, Speed Up, Slow Down
Touch	Touch, Push
Verbal	To Agent, About Agent, Spontaneous Utterance, Laughing
Operational	Wire Adjustment, Wire Management, Ask Researcher
Other	Special Note

Table 9: Behaviour codes. Descriptions/ examples for each can be found in the appendices

The codes from each video were counted and the mean number of occurrences for each code for each condition was computed. The Friedman test was used, and eight codes were identified as showing statistically significant ($p < 0.05$) difference between conditions, shown in *Table 10*. Wilcoxon signed-rank post-hoc to identify statistically significant differences between pairs of conditions within the identified behaviours ($p < 0.017$, Bonferroni correction), shown in *Table 11*.

Index	Code	χ^2 (2)	p
1	Change direction	10.932	0.004
2	Slow down	7.752	0.021
3	Stop	13.303	0.001
4	Watching	12.521	0.002
5	Avoidance Gesture	7.389	0.025
6	Minor Reaction	8.6	0.014
7	Protective Gesture	8.087	0.018
8	Sudden Stop	21.726	0.0005

Table 10: Statistically significant behaviours

Index	Conditions A/B	μ_A	μ_B	Z	p
Change Direction	Low-Med	1.86	3.09	-2.835	0.005
	Low-High	1.86	4.0	-3.216	0.01
Stop	Low-High	1.14	2.77	-3.370	0.001

	Med-High	1.45	2.77	-2.668	0.008
Watching	Low-High	1.95	0.55	-3.25	0.001
	Med-High	1.77	0.55	-3.00	0.003
Avoidance Gesture	Low-High	0.14	0.91	-2.388	0.017
Minor Reaction	Med-High	0.05	1.86	-2.987	0.004
Sudden Stop	Low-High	0.86	2.41	-3.774	0.0
	Med-High	1.09	2.41	-2.974	0.003

Table 11: Post-hoc tests on significant behaviours.

The codes in *Tables 10* and *11* mainly relate to actions over a distance and reactive behaviours (showing the largest difference in occurrence). Observational behaviours occurred at high frequencies in all conditions, but show no significant differences between densities except for the watching behaviour which was more prevalent in the low and medium densities. As with the affective state measure, the highest differential was between High and Low than the other two

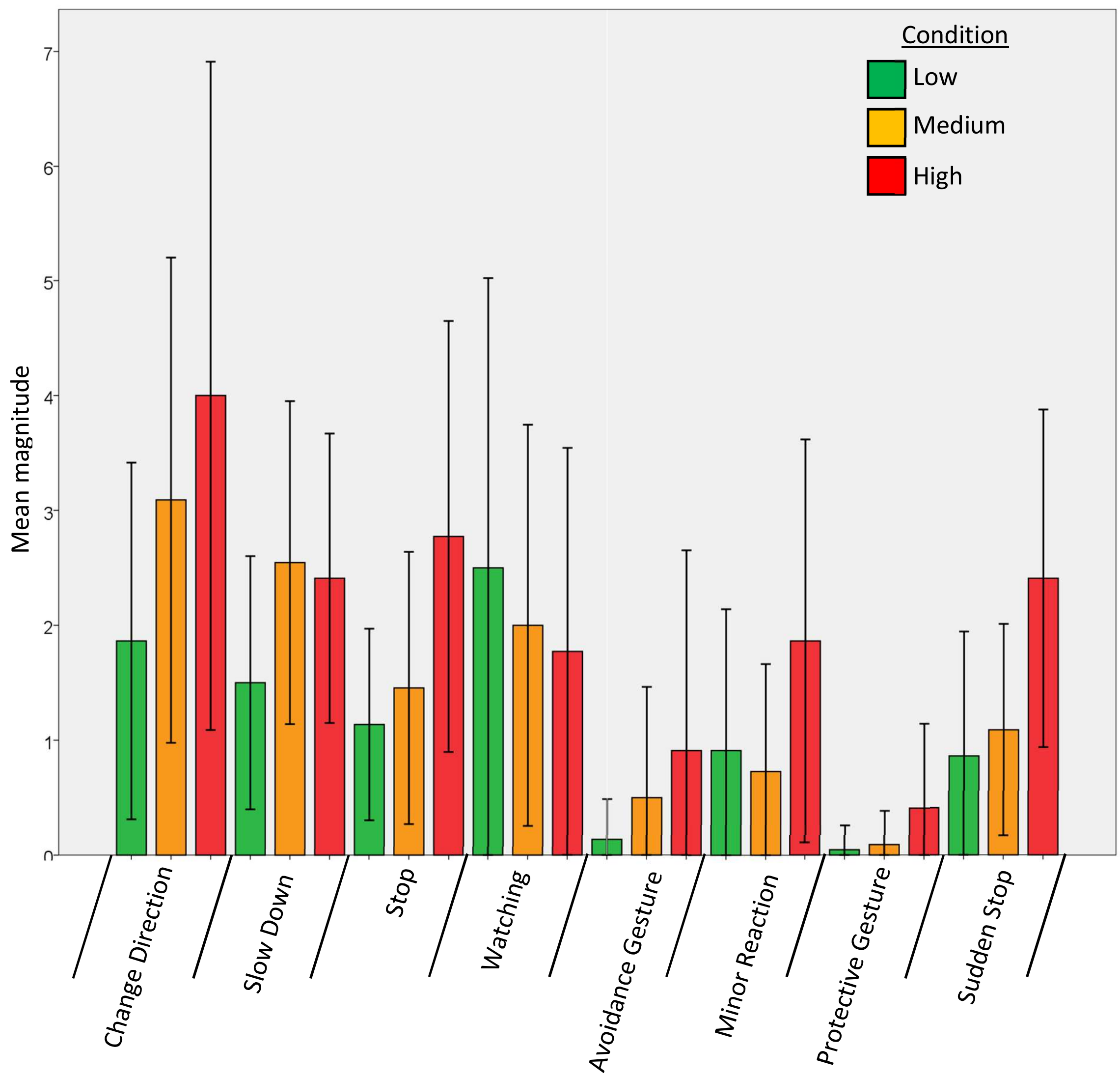


Table 12: Graph showing mean occurrences of significant behaviour across the 3 conditions.

comparisons. A table showing the sum of each code between condition conditions can be found in *Appendix 7*.

5.6.4 Interviews

The interviews were not thematically analysed, but they did provide useful qualitative data from each participant. Questions ranging from how they feel about real crowds to how they felt throughout the study and how difficult they found the task between different conditions. More details on each question can be found in *Section 3.2.3.* Excerpts from these interviews were used to supports points raised in the discussion segment, *Section 5.6.6*.

5.6.5 Virtual Participant Trajectory Analysis

As previously mentioned, it was decided that getting an agent to (simulate) completing the task would help show any differences between them and the real participants. Such things as average movement speed or average proximity to other agents could be compared. The average proximity between the simulated participant and other agents was calculated the same way as with the real participants, as explained earlier in *Section 5.6.2*. A repeated measures ANOVA (with Greenhouse-Geisser correction) showed significant differences between conditions (F (1.633,37.57) = 174.227, $p < 0.0005$). Post-hoc with Bonferroni correction showed significant differences between all conditions ($p < 0.05$). It is noted that on a comparison between the real participant and virtual for high density, the average proximity shows a reasonably small difference in magnitude but are still significantly different at $p < 0.01$ (using a 2-tailed t-test). The implications of this will be discussed in next section.

Condition	Distance to nearest agent (m)	
	Mean (μ)	Standard Deviation (σ)
Low	2.13	0.30
Medium	1.37	0.22
High	0.98	0.07

Table 13: statistical summary for the virtual participant's proximity to other agents

5.6.6 Findings

5.6.6.1 Response to Agent Density

The overall reaction to the experience within the simulation from participants was positive with descriptions including "fun" and or "enjoyable". In some cases, participants reported this being because they had never used VR before, although more than half had done so in the past and ten of those have used the Vive. The consensus from participants was that the simulation shows aspects of realism, although it had limitations. One participant is quoted as saying, "*I thought the people were real although they behaved unusually. But it felt like a realistic experience.*"

The significant increase in negative affect is the key result of this project. Comments such as "*The more people there were the more frustrating I think it was*", and "*it got more claustrophobic with more people*", another "*I felt more anxious and more distressed because there were lots of people in the way*" reflects this. Words such as "anxiety" "frustrating" and "nervous" were often used when describing some aspects of the system. Participants were asked how they felt about crowds in real life to which a majority reported their dislike for real crowds, and mentioned feeling similar feelings and reaction during the study.

Analysis showed that on average the participants were closer to agents during higher density conditions, which is not surprising. Trajectory data showed the mean separation between the user and the closest agent decreases from 2.08m to 1.05m (*Table 8*), a difference of 1.06m. This shows a clear change in interpersonal distance, with it changing from a social distance of 1.2m – 3.6m to within one's personal space of 1.2m or less. Wilcox et al (2006) for example incremented the distance by only 0.5m and found a clear difference in participant response between conditions. This is further supported by observational data.

Outlined in *Section 5.6.3*, some participant reactions were categorised as "reactive", to indicate a sudden reaction to agents, one which may not have had any conscious thought such as a surprised jump. These behaviours were more prevalent in high density, as opposed to low and medium. Of the eight significant behaviour changes, four of them were in this category. The proximity may not have been the only influencer however with participants possibly becoming overwhelmed by the sheer number of agents rather than their proximity. One participant, for example, remarked, "*when there were less people it seemed less worrying cause there was more space and you could see a line to go through*". During high density one participant had to be prompted on what the task was even after completing it twice before; another commenting "*it was the only time I forgot what I was supposed to do*". Showing a clear sign that the agents were more distracting at higher densities.

5.6.6.2 Response to Agent Behaviour

Observational and interview data revealed some points regarding the participants' interpretation and reaction to the agents' individual behaviours. Two questions were asked relating to this.

- To what extent do you think that the agent's behaviour was realistic was it like being in a real crowd of people?
- Do you feel that the agents were aware of you, and responded to you?

Participants responses showed some initial uncertainty, leading to the inclination to experiment with agent responsiveness. For example, "*I wasn't sure what to expect because they weren't human*"; and "*I felt more confident as I got used to it, because I knew what to expect from the movement of people*". Nine instances of a participant trying to touch an agent were recorded, and seven times when a participant talked directly to an agent. To some extent participants did try to interact with agents in a realistic way. Some went into how they felt about this with quotes such as "*the feeling of not bumping into people felt a bit strange at times*"; and "*I prodded one of them to see what happened, my hand went through, so I thought that was cool. I felt like he was a ghost*".

Regarding the realism of the agents, participants gave mixed responses. Several described the agents as being "rude" or "aggressive"; for example, "*they didn't stop and give way as you would expect from a real crowd... they tended to be more aggressive*". Another commented, "*if they saw me coming they wouldn't change direction, it would almost elevate hostility. You almost feel like you're being walked over*". Similarly: "*I'd like to think people would stop for you in a real crowd*", "*They seemed like they were in a rush.*" Crowd simulations, in general, are designed to reproduce realistic behaviour (given scalability constraints) but are not and do no capture the complexity of human interactions in crowded situations. Whilst this facilitates the concept of validation through observation (e.g. Pelechano et al, 2008b), these comments suggest that they interpret agent behaviour in a negative light (from lack of interaction) possibly contributing the negative affect observed from responses.

Response regarding agent awareness gave similar feelings, and suggest participants were frustrated at being unable to interact with agents in a natural way. One participant is quoted "*I felt they were aware of me when they were in proximity to me. I didn't feel they were aware of me in their trajectory*", another "*They seemed like they had semi-awareness*". Another commented, "*they weren't stopping as they would in real life, like stop and let me go past, obviously, but they were changing direction if I was there.*" And another similarly noted "*I think they were more determined to walk in a straight line than I was, so I felt like I was the only one trying to avoid them. But they were quite responsive if they did get really close.*" These two comments correctly describe the social forces model used, with less movement impact far away compare to up close, showing that participants could assess the behaviour naturally; but were less able to engage with it.

By further comparing the average proximity data from the virtual participant (*Table 13*) and real participants (T*able 8*) a small but significant difference becomes apparent. While it is a small difference (0.07m in high density) it may indicate a natural tendency or agents to get closer to each other than what the human participants found comfortable, possible contributing to feelings of discomfort and intimidation further raising the negative affect felt at higher densities.

5.6.6.3 Participant Behaviour

In the previous section, it was mentioned that from observational analysis and interviews that participants experience a degree of experimentation and learning. They readily engaged with this without prompting from the Researcher. Previous work (Narang et al, 2016; Pelechano et al, 2008b) have reported similar observations. However, these behaviours are relativity short-lived and occur infrequently once the participant has assessed the agent's range of behaviours. Two questions were asked regarding their own behaviour:

- *Did you try and interact with the agents?*
- *Apart from the restrictions of the Vive, did you find it difficult to move around the environment (more difficult than in a real crowd)?*

Some participants commented on how they tried to speak with or physically touch agents. For example, "*I did say sorry once which was a bit odd.*" The limited ability in which the participants could interact with the agents also came up, with participant adjusting their behaviour accordingly. For example, one participant states "*In real life I would apologise if I bump into people... I never said anything.*" Another "*The crowd's activity didn't really affect me at all, as soon as I realised that me bumping into them had no effect it didn't seem to matter.*" There is evidence to suggest that once the participants realised that the agents would not engage in a realistic way they felt they had to be more accommodating. For example, "*I just waited for longer for the people to move. I don't know how people normally move through crowds but I would wait till there is a gap before moving forward.*"

Some participants seemed to be able to disassociate themselves from the simulation to complete the task more efficiently. One commented "*you realise they just bounce off of you after a while... oh well, just carry on walking through them*" and another held "*The crowd's activity didn't really affect me at all, as soon as I realised that me bumping into them had no effect it didn't seem to matter*"

The video annotations also highlighted the significant difference between the conditions. Tables and detailed statistically significant behaviours, behaviour were classified into groups as either "action" or "reaction", where "action" behaviours were to be considered responses to agent behaviours (such as waiting for an agent to pass), and "reaction" to be sudden reactions (such as raising hands or stepping away). It is reasonable to expect an increase in both types with increasing density, and indeed the total number of action behaviours rose from 160 at low density (summed across all participants) to 245 at high density. However, there was a much larger rise in reaction codes, from 60 to 158. This suggests that participants were less able to plan interactions with agents and instead had to respond reactively to agents (possibly because they were closer or just more numerous). It might be reasonable to expect that reactive behaviours indicate higher levels of discomfort or anxiety, and so be associated with the higher levels negative affect seen in higher density situations.

6 Discussion

This project aimed to complement and extend existing crowd simulation work in VR by exploring the impact crowd density has on the user's experience; by measuring the user's affective state while performing a simple task-based activity in a continuous social-forces based model. Results showed a significant increase in negative affect in high-density crowd situations along with significant increase in proximity, a decrease in movement speed and changes in some behaviours.

This section will discuss the impact these results have with the focus on the effects of crowd density in VR; the implications for new crowd behaviour models and applications for VR crowds. The findings are also generalised beyond the experimental setting to…

6.1 Effects of Crowd Density Within VR

One key motivation for this project was the ability to embed a human user in an existing simulation model, to develop a greater understanding for how people interact with crowds in general and of varying densities, possibly leading to more sophisticated outcomes.

The main advantage of incorporating a human participant into the simulation is that the agents' models are relatively simple and cannot capture the full range of human interactions and behaviour while in a crowd (e.g. Moussaïd et al, 2016). Density is an important factor when dealing with large groups of people, for example situations such as evacuations and festivals have very different crowd behaviour depending on size. The results from this study show a clear response to change in density, it can thus be asserted that the simulation of such scenarios could help better inform processes such as building design, evacuation planning and event management.

Previous work into proxemics in VR used either a stationary participant or non-natural control methods such as hand help controllers. The study used the HTC Vive platform, which allowed users to walk around the world in a natural way (even if restricted to a 2mx2m area), they could also interact with an object with two tracked hand controllers. The kinaesthetic dimension allowed participant's to have finer control over objects, proximity and behaviour which were expressed naturally through body movement. This also meant the participant's speed, balance and geometry were naturally utilised which is difficult to do with any other form of input.

6.2 Using Self-Reported Affective State As A Measure For Human Response To Crowd Density

This study is the first to use the PANAS questionnaire in the relation to user affect while interacting with a VR crowd simulation, previous work has used participant presence (Whitmer and

Singer, 1998), bespoke questionnaires, trajectory data, and physiological measures such a GSR. While these are revealing and useful metrics in themselves they do not cover the entire spectrum of user experiences. Earlier work linked presence and emotional response (Diemer et al, 2015); affect was used within this research due to its ability to give a more direct and rounded measure of user experience from a design point of view. Participant behaviour has also been systematically categorised and quantified, this measure may provide a useful additional tool for future study.

6.3 Reconsidering Crowd Behaviour Models

Opportunities arise when both qualitative and quantitative data measures are used. Greater insight into the quantitative with qualitative data slows greater insight into different crowd behaviour models further addressing participant response to the simulation and agent behaviour.

Negative comments were made relating to agent behaviour by some participants, descriptors such as "aggressive or "rude" indicate a negative response or interpretation. These relate to the simulation directly, which like most crowd simulations are not designed explicitly for direct human-agent interaction. This raises many questions into the design of simulations which incorporate human participant's, this warrants further study and evaluation.

Validation of a crowd simulations usually falls back on its macroscopic behaviour (Fundamental diagram for example (*Section 4.2.3*; Seyfried, et al, 2008)) rather than microscopic interactions, this could lead to tension between the need to create larger/ scalable models and the realistic interactions needed on an individual level for a user to interact with it. Previous work into validation through presence does little to explore this as the user is cast as an observer rather than an active participant.

Some studies have considered the using gestures and other subtle interactions in simulations, by layering these embellishments on top of existing models, which may go some way to alleviate these issues. This had led to some interesting results from studies such as Narang et al (2016) and Kyriakou et al (2015); however, implementing these for varying simulation models and conditions is complex. For example, conflicting and confusing behaviours may arise if positive embellishments become negative under different scenarios. For example, an agent making eye contact and smiling at the user could be positive and realistic in a sparsely populated simulation but if all the agents did the same in a high-density situation it could become intimidating and/or upsetting.

6.4 Application for Crowd Simulations Within VR

The application for VR crowd simulations have been alluded to previously in this paper, including training, event planning and evacuations. Given the results show a significant impact on behaviour and emotional state from the virtual crowd density, study into simulation preparing individuals for challenging scenarios involving crowds might be beneficial. An example such as guiding an evacuation, crowd management at events. Similarly, VR simulations do not only have to be used to induce a negative effect from the large crowds but to also offer the opportunity to study environmental factors that contribute to the difficulty of different behaviours such as lighting/ visibility, environmental noise, etc.

Discussed previous, VR is widely used in therapy to treat many phobias and anxiety causing conditions (Opris et al, 2012). Further study should be done into the use of VR crowd simulations into therapy for conditions relating to the fear of crowds (agoraphobia) or situations with heightened anxiety levels related to crowds.

The Social forces model has not only been used for crowd simulations but also in robotic for example, such models are used to mediate human-robot interactions (HRI; e.g. Ferrer et al, 2013). So, the use of VR crowd simulation could, therefore, be as an effective HRI research tool to test prototype force models and enabling the creation of extraordinary scenarios.

7 Conclusion

This project has achieved its goal of evaluating the change in user response when interacting with an immersive VR crowd simulation of varying crowd density. The change in behaviour has also been assessed along with a comparison made between the participant behaviour and a simulated agents behaviour whilst performing the same task. This section will summarise all of this and give insight into its limitations and discusses possible areas of study for future work.

Comparisons between human and agent behaviour showed significant differences in even with the use of a validated model. This presents a dilemma for the credibility of current validation methods used such as the speed vs density fundamental diagram. Some participants described specific unrealistic behaviours from the agents such as collisions and their reluctance to stop for other agents/people. This raises the belief that the use of an egocentric perspective within VR can offer greater insight into model validity than other methods.

The results gathered demonstrate a significant change in emotional response and behaviour whilst interacting in the crowd simulation. The negative affect increased significantly at higher densities correlating with other studies into proxemics with and without VR. Along with this, instinctive behaviours, such as startled jumps and protective gestures also increased in users. Considered actions such as slowing for agents at a distance and changing trajectory to walk around them increased significantly with density also. These behaviour changes correlate with those from real crowd situation. This opens new opportunities to study human behaviour under many simulated crowd situations, it also suggests that VR can be leveraged to study how people would react under similar real-situations.

7.1 Future Work and Limitations

This project has considered crowd simulation in VR in a general sense by implementing a full simulation with user interaction, it has also looked at the effect density has on the participants' affective state and behaviour. By using multiple measures, including behavioural observation, interviews and an affective state questionnaire, it has given greater insight into how effective this combination of technology could be. Nevertheless, the project has some limitations which are discussed below.

Participants interacted with the simulation for short periods of time, with each scenario lasting three to four minutes each. To make more accurate comparison between user reaction within real life and virtual crowds situations, a longer study would be necessary. Real life crowd situations would usually be experienced for longer periods of time than was possible in this study.

Arguably this increased time within the virtual crowded setting would raise new, yet unobserved emotions and behaviours. Similarly, some of the participants had never experienced using a VR device prior to taking part, which is likely to have had an impact on their emotional state while participating. A longer-term study would be beneficial, allowing time for participants to become familiar with the simulation and VR.

PANAS allows for an effective consideration of the users' affective state at the time of completing the Questionnaire, it has been validated by multiple studies as accurate when compared to physiological data (*Section 3.2.1*). However, the participants had to leave the simulation and VR to complete the form, taking anywhere from thirty seconds to 2 minutes before they could begin. This transition time may have affected their response. Arguably if the limitations of physiological tracking, discussed in *Section 3.2.1*, were effectively managed it would allow for the timely collection of objective data without a break in immersion for the participants. The data gathered could also be connected to specific events that happen within the simulation, this together with the use of PANAS would arguably provide a rich picture; showing the situation specific physiological affective state for the whole scenario in addition to the subjective perspective of the participants.

By comparing participant behaviour to that of simulated agents completing the same task, it was found that the simulated agents spent significantly more time closer to other agents at higher densities (*Sections 5.6.2* and *5.6.5*). This shows that even though care was taken to model human behaviour in a realistic way, discrepancies are still present, more work needs to be done in this area, perfect simulation of human behaviour is likely to be a future challenge.

As previously discussed (*Section 3.1.1*) the VR technology used, HTC Vive, has a restricted walkable area, typical of any room scale VR system. This limits the possible scenarios for which this technology can be used. Other methods of locomotion within scenarios are possible however, and by using one of these or a combination of them, the restrictions could be minimised or perhaps alleviated. Possible techniques such a teleportation have been used extensively in video games allowing the user to move around the virtual world, and so avoiding actual real-life movement over large distance outside of the tracked area. However, this may cause other problems due to the lack of realism, which was the main reason they were not used in this study.

The limited field of view available to the user was shown to be a drawback, as outlined in the introduction (S*ection 1*). Multiple participants commented on how they were not able to see through their periphery so had to turn their heads more. This change in behaviour due to technological limitations is not ideal. The wired connection of the Vive also caused some breaks in immersion for the participants, as they had to sometimes stop to manage the wire, by using a wireless system this could be avoided.

Despite these difficulties it seems clear that VR crowd simulation has the potential to be an invaluable resource for future development.

9 References

Achibet, M. Girard, A. Marchal, M. Lécuyer, A. (2016) Leveraging passive haptic feedback in virtual environments with the elastic-arm approach. Presence: Teleoperators and Virtual Environments, 25(1), pp.17-32.

Adobe Systems (2014) *Fuse Character Creator* [software] 1.2. CA, USA: Adobe. Available from http://www.adobe.com/uk/products/fuse.html [accessed 7th September 2017]

Afnaan, S. (2017) An introduction to Virtual Reality (VR). *Technobyte,* 30th May 2017.

Ahlberg, G. Enochsson, L. Gallagher, A.G. Hedman, L. Hogman, C. McClusky, D.A. Ramel, S. Smith, C.D. Arvidsson, D. (2007) Proficiency-based virtual reality training significantly reduces the error rate for residents during their first 10 laparoscopic cholecystectomies. *The American journal of surgery*, 193(6), 797-804.

Ahn, J. Wang, N. Thalmann, D. Boulic, R. (2012) Within-crowd immersive evaluation of collision avoidance behaviors. In: *Proceedings of the 11th ACM SIGGRAPH International Conference on Virtual-Reality Continuum and its Applications in Industry*. ACM. 231-238

Alaraj, A. Luciano, C.J. Bailey, D.P. Elsenousi, A. Roitberg, B.Z. Bernardo, A., Banerjee, P.P. Charbel, F.T. (2015) Virtual reality cerebral aneurysm clipping simulation with real-time haptic feedback. *Operative Neurosurgery*, *11*(1), pp.52-58.

BBC News (1989) Football fans crushed at Hillsborough. *BBC News,* 15th April 1989. Available from http://news.bbc.co.uk/onthisday/hi/dates/stories/april/15/newsid_2491000/2491195.stm [accessed 9th September 2017]

Braun, A. Musse, S. de Oliveira, L. Bodmann, B. [2003] Modelling Individual Behaviours in Crowd Simulation. *IEEE,* 2003.

Bruneau, J. Olivier, A.H. Pettre, J. (2015) Going through, going around: A study on individual avoidance of groups. IEEE transactions on visualization and computer graphics, 21(4). 520-528.

Burstedde, C. Klauck, K. Schadschneider, A. Zittartz, J. (2001) Simulation of Pedestrian Dynamics Using a Two-Dimensional Cellular Automaton. *Physica A: Statistical Mechanics and its Applications*, 295(3), 507-525.

Buttussi, F. and Chittaro, L. (2017) Effects of different types of virtual reality display on presence and learning in a safety training scenario. *IEEE transactions on visualization and computer graphics*.

Chen, D. Wang, L. Wu, X. Chen, J. Khan, S.U. KołOdziej, J. Tian, M. Huang, F. Liu W. (2013) Hybrid modelling and simulation of huge crowd over a hierarchical grid architecture. *Future Generation Computer Systems*, 29(5) pp.1309-1317.

Christou, C. Herakleous, K. Tzanavari, A. and Poullis, C. (2015) Psychophysiological responses to virtual crowds: Implications for wearable computing. In *Affective Computing and Intelligent Interaction (ACII), 2015 International Conference*. IEEE. 35-41.

Crawford, J. Henry, J. (2004) The Positive and Negative Affect Schedule (PANAS): Construct validity, measurement properties and normative data in a large non-clinical sample. *Clinical Psychology,* 43(3), 245-265.

Dictionary.com (2017) http://www.dictionary.com/ [accessed 18th August 2017]

Diemer, J. Alpers, G.W. Peperkorn, H.M. Shiban, Y Mühlberger, A. (2015) The impact of perception and presence on emotional reactions: a review of research in virtual reality. Frontiers in psychology, 6.

Durupinar, F. Allbeck, J. Pelechano, N. Badler, N. (2008) Creating crowd variation with the ocean personality model. *In Proceedings of the 7th international joint conference on Autonomous Agents and Multiagent Systems (AAMAS 2008)*, Padgham, Parkes, Müller and Parsons (eds.). 1217-1220

Egan, D. Brennan, S. Barret, J. Qiao, Y. Timmerer, C. Murray, N. (2016) An evaluation of Heart Rate and Electro Dermal Activity as an objective QoE evaluation method for immersive virtual reality environments. *2016 Eighth International Conference on Quality of Multimedia Experience (QoMEX)*, Lisbon, Portugal. IEEE. 1-6.

Epic Games (2015) *Unreal Engine*. [software] 4.15. Cary, NC: Epic Games. Available from https://www.unrealengine.com/what-is-unreal-engine-4. [accessed 28th October 2017]

Ferrer, G. Garrell, A. Sanfeliu, A. (2013) Robot companion: A social-force based approach with human awareness-navigation in crowded environments. In*: Intelligent robots and systems (IROS),* 2013 IEEE/RSJ international conference. IEEE. 1688-1694.

Frontier Developments (2016) *Planet Coaster* [Video Game]. Cambridge: Frontier Developments. Available from https://www.planetcoaster.com/en [accessed 14th August 2017]

Garau, M. Slater, M. Pertaub, D.P. Razzaque, S. (2005) The responses of people to virtual humans in an immersive virtual environment. *Presence: Teleoperators and Virtual Environments*. 14(1). 104-116.

Golas, A. Narain, R. Curtis, S. Lin, M.C. (2014) Hybrid Long-Range Collision Avoidance for Crowd Simulation. *IEEE transactions on visualization and computer graphics*, 20(7) 1022-1034.

Gugenheimer, J. Wolf, D. Eiriksson, E.R. Maes, P. Rukzio, E. (2016). Gyrovr: Simulating inertia in virtual reality using head worn flywheels. In Proceedings of the 29th Annual Symposium on User Interface Software and Technology (pp. 227-232). ACM. Vancouver

Hall, E. (1966) *The hidden Dimension*. NY, USA: Anchor Books.

Hall, E (1968) Proxemics. *Current Anthropology* 9(2-3) 83-108.

Helbing, D. Molnár, P (1995) Social Force Model for Pedestrian Dynamics. *Physical Review E,* 51(5) 4282-4286.

Hupont, I., Gracia, J., Sanagustín, L., & Gracia, M. A. (2015). How Do New Visual Immersive Systems Influence Gaming QoE? A use case of serious gaming with Oculus Rift. *Quality of Multimedia Experience (QoMEX), 2015 Seventh International Workshop on,* 26 – 29 May 2015, 2 – 6.

Jennett, C. Cox, A.L. Cairns, P. Dhoparee, S. Epps, A. Tijs, T. Walton, A. (2008) Measuring and defining the experience of immersion in games. *International journal of human-computer studies*, 66(9), Elsevier. 641-661.

Jiang, H. Xu, W. Mao T. Li, C. Xia, S. Wang, Z., (2009) A semantic environment model for crowd simulation in multi-layered complex environment. In *Proceedings of the 16th ACM Symposium on Virtual Reality Software and Technology*. ACM. 191-198

Johansson, A. Helbing, D. Shukla, P.K. (2007) Specification of the social force pedestrian model by evolutionary adjustment to video tracking data. *Advances in complex systems*, *10*(supp02). 271-288.

Jung, T. Tom Dieck, M.C. Lee, H. Chung, N. (2016) Effects of virtual reality and augmented reality on visitor experiences in museum. In Information and Communication Technologies in Tourism 2016 (pp. 621-635). Springer, Cham.

Kastanis, I. and Slater, M. (2012) Reinforcement learning utilizes proxemics: An avatar learns to manipulate the position of people in immersive virtual reality. *ACM Transactions on Applied Perception (TAP)*. 9(1). ACM. 3.

Kim, S. Bera, A. Best, A. Chabra, R. Manocha, D. (2016). Interactive and adaptive data-driven crowd simulation. In *Virtual Reality (VR), 2016 IEEE*. IEEE. 29-38

Kyriakou, M. Pan, X. Chrysanthou, Y. (2015) Interaction with Virtual Agents—Comparison of The Participants' Experience Between An IVR And A Semi-IVR System. In Virtual Reality (VR), 2015 IEEE. IEEE. 217-218.

Kyriakou, M. Pan, X. Chrysanthou, Y. (2016) Interaction with virtual crowd in Immersive and semi-Immersive Virtual Reality systems. *Computer Animation and Virtual Worlds 28 (5)*. John Wiley & Sons Ltd.

Laerde Statistics (2013) Sphericity. Laerd Statistics. Available from https://statistics.laerd.com/statistical-guides/sphericity-statistical-guide.php [accessed on 6th February 2018]

Laffont, P.Y. Martin, T. Gross, M. Tan, W.D. Lim, C.T. Au, A. and Wong, R. (2016) Rectifeye: a vision-correcting system for virtual reality. In *SIGGRAPH ASIA 2016 Posters* (p. 8). ACM.

Lemercier, S. Jelic, A. Kulpa, R. Hua, J. Fehrenbach, J. Degond, P. Appert-Rolland, C. Donikian, S. Pettré, J. (2012) Realistic following behaviors for crowd simulation. In: *Computer Graphics Forum*. Blackwell Publishing Ltd. 31(2.2). 489-498

Llobera, J. Spanlang, B. Ruffini, G. and Slater, M. (2010) Proxemics with Multiple Dynamic Characters in An Immersive Virtual Environment. *ACM Transactions on Applied Perception (TAP)*, *8*(1). 3-21.

Meehan, M. (2001*) Physiological reaction as an objective measure of presence in virtual environments*, PhD. University of North Carolina at Chapel Hill. Available from http://www.cs.unc.edu/techreports/01-018.pdf [accessed 10th October 2017]

Mehran, R. Oyama, A. Shah, M. (2009) Abnormal crowd behavior detection using social force model. In: *Computer Vision and Pattern Recognition, 2009*. CVPR 2009. IEEE Conference. IEEE. 935-942

Microsoft Studios (2011) *Minecraft* [Video Game] WA, USA: Microsoft Studios.

Moussaïd, M, et. al. (2016) Crowd behaviour during high-stress evacuations in an immersive virtual environment. *Journal of the Royal Society Interface* 13(122).

Morgan, D.L. (1998) Practical strategies for combining qualitative and quantitative methods: Applications to health research. *Qualitative health research*, 8(3), pp.362-376.

Neira, C. Sandin, D. DeFanti, T. (1993) Surround-screen Projection-based Virtual Reality: The Design and Implementation of the CAVE. In: *SIGGRAPH '93 Proceedings of the 20th annual conference on Computer graphics and interactive techniques,* Anaheim, CA, 2nd – 6th August 1993. New York, USA: ACM, 135 – 142.

Narang, S. Best, A. Randhavane, T. Shapiro, A. and Manocha, D. (2016) PedVR: Simulating gaze-based interactions between a real user and virtual crowds. In *Proceedings of the 22nd ACM conference on virtual reality software and technology*. ACM. 91-100.

North, M. North, S. (2016) Virtual Reality Therapy. In: *Computer-Assisted and Web-Based innovations in Psychology, Special Education, and Health*. Luiselli, J. K. Fischer, A. J. (ed.) London: Elsevier Inc, 141-156.

O'Brolcháin, F. Jacquemard, T. Monaghan, D. O'Connor, N. Novitzky, P. Gordijn, B. (2016) The convergence of virtual reality and social networks: threats to privacy and autonomy. *Science and engineering ethics*, *22*(1), pp.1-29.

Olivier, A.H. Bruneau, J. Cirio, G. Pettré, J. (2014) A virtual reality platform to study crowd behaviors. *Transportation Research Procedia 2*. Elsevier. pp.114-122.

Opriş, D. Pintea, S. García-Palacios, A. Botella, C. Szamosközi, Ş. David, D. (2012). Virtual reality exposure therapy in anxiety disorders: a quantitative meta-analysis. *Depression and anxiety*, 29(2). 85-93.

Passos, E. Joselli, M. Zamith, M. Rocha, J. Montenegro, A. Clua, E. Conci, A. Feijo, B. (2008) Supermassive Crowd Simulation on GPU Based on Emergent Behavior. In: Proceedings of the VII Brazilian Symposium on Computer Games and Digital Entertainment (pp. 81-86).

Pesce, M (2015) *HTC executive director of marketing, Jeff Gattis wears the HTC Re Vive.* [image] available from https://www.flickr.com/photos/pestoverde/16513847554 [accessed 7th August 2017]

Pelechano, N. Allbeck, J.M. (2016). Feeling crowded yet?: crowd simulations for VR. In *Virtual Humans and Crowds for Immersive Environments (VHCIE), IEEE*. IEEE. 17-21.

Pelechano, N. Allbeck, J. Badler, N. (2007a) Controlling Individual Agents in High-Density Crowd Simulation. *Proceedings of the 2007 ACM SIGGRAPH/Eurographics Symposium on Computer Animation*, 99-108.

Pelechano, N. Allbeck, J. Badler, N. (2008a) *Virtual Crowds: Methods, Simulation, and Control.* Synthesis Lectures on Computer Graphics and Animation, 3(1), 1-176.

Pelechano, N. Stocker, C. Allbeck, J. Badler, N. (2007b) feeling Crowded Yeat? Exploring Presence in Virtual Crowds. In: Annual International Workshop on Presence "proceedings of Presence 2007". 373-376.

Pelechano, N. Stocker, C. Allbeck, J. Badler, N (2008b) Being a Part of the Crowd: Towards Validating VR Crowds Using Presence. *Proc. Of 7th Int. Conf. on Autonomous Agents and Multiagent Systems (AAMAS 2008)*. Padgham, Parkes, Müller and Parsons (eds.), 12 - 16 May 2008.

Qiu, F. and Hu, X. (2010). Modeling group structures in pedestrian crowd simulation. *Simulation Modelling Practice and Theory*. 18(2). Elsevier. 190-205.

Red Storm Entertainment (2017) *Star Trek: The Bridge Crew* [Video Game] Rennes, France: Ubisoft.

Reynolds, C.W. (1987) Flocks, herds and schools: A distributed behavioral model. ACM SIGGRAPH computer graphics, 21(4), 25-34.

Rojas, F.A. Yang, H.S. (2013) Immersive Human-In-The-Loop Hmd Evaluation of Dynamic Group Behavior In A Pedestrian Crowd Simulation That Uses Group Agent-Based Steering. In Proceedings of the 12th ACM SIGGRAPH International Conference on Virtual-Reality Continuum and Its Applications in Industry. ACM. 31-40.

Rojas, F.A. Yang, H.S. Tarnogol, F.M. (2014) Safe Navigation of Pedestrians in Social Groups in a Virtual Urban Environment. In: *Cyberworlds (CW), 2014 International Conference.* IEEE. 31-38.

Russell, W.D. Newton, M. (2008) Short-term psychological effects of interactive video game technology exercise on mood and attention. Journal of educational technology & society, 11(2).

Sanchez-Vives, M.V., Slater, M. (2005). From Presence to Consciousness Through Virtual Reality. In: *Nature Reviews Neuroscience*, 6(4), 332-339.

Sarkar, P. (2000) A Brief History of Cellular Automata. ACM Computing Surveys (csur), 32(1), 80-107.

Schadschneider, A. Seyfried. A. (2009) Validation of ca models of pedestrian dynamics with fundamental diagrams. *Cybernetics and Systems*, 40(5): 367–389.

Schubert, T. Friedmann, F. Regenbrecht, H. (2001) The experience of presence: Factor analytic insights. *Presence: Teleoperators and virtual environments*, 10(3). 266-281.

Schumm, J. Bachlin, M. Setz, C. Arnrich, B. Roggen, D. and Troster, G. (2008). Effect of movements on the electrodermal response after a startle event. In *Pervasive Computing Technologies for Healthcare, 2008. PervasiveHealth 2008. Second International Conference*. IEEE. 315-318.

Seitz, M. Deitrich, F. Köster (2014) A Study of pedestrian Stepping Behaviour for Crowd Simulation. In: *The Conference on Pedestrian and Evacuation Dynamics 2014 (PED2014),* Delft, Netherlands, 22 – 24 October 2014. Netherlands: Elsvier, 282 – 290.

Seyfried, A. Steffen, B. Klingsch, W. Boltes, M. (2005) The Fundamental Diagram for Pedestrian Movement Revisited. In: *Journal of Statistical Mechanics: Theory and Experiment,* 2005(10), 10002.

Shao, W. Terzopoulos, D. (2005) Autonomous pedestrians. In: *Proceedings of the 2005 ACM SIGGRAPH/Eurographics symposium on Computer animation*. ACM. 19-28

Slater, M. (2002) Presence and the sixth sense. Presence: Teleoperators and Virtual Environments, 11(4), pp.435-439.

Slater, M. (2003) A Note on Presence Terminology. *University College London.* Available from http://www0.cs.ucl.ac.uk/research/vr/Projects/Presencia/ConsortiumPublications/ucl_cs_papers/presence-terminology.htm [Accessed 10th August 2017]

Sohre, N. Mackin, C. Interrante, Guy, S. (2017) Evaluating Collision Avoidance Effects on Discomfort in Virtual Environments. In: *Virtual Humans and Crowds for Immersive Environments (VHCIE),* LA, USA, 19th March 2017. IEEE. Virtual Humans and Crowds for Immersive Environments (VHCIE), 2017 IEEE.

Steffen, B. Seyfried, A. (2008) The repulsive force in continous space models of pedestrian movement. NY, USA: arXiv.

Sugand, K. Akhtar, K. Khatri, C. Cobb, J. Gupte, C. (2015) Training effect of a virtual reality haptics-enabled dynamic hip screw simulator: A randomized controlled trial. *Acta orthopaedica*, *86*(6), pp.695-701.

Taylor, J. (1990) *Hillsborough Disaster Working Party.* London, UK: Her Majesty's Stationary Office. Available from http://hillsborough.independent.gov.uk/repository/HWP000000180001.html [accessed 11th October 2017]

Teslasuit (2017) Tesla Suit. Available from https://teslasuit.io/ [accessed 22nd November 2017]

Thomas, J. Bashyal, R. Goldstein, S. Suma, E. (2014) MuVR: A multi-user virtual reality platform. In Virtual Reality (VR), 2014 IEEE (pp. 115-116). IEEE.

Valve Corporation (2016) *The Lab* [Video Game] Washington, USA: Valve Corporation

Ubisoft Montreal (2014) *Assassin's Creed: Unity*. [Video Game] Carentoir, France : Ubisoft.

Unity Technologies (2016) *Unity3D*. [software] 5.3.3. San Fransisco, CA: Unity Technologies. Available from http://unity3d.com/. [accessed 28th August 2017].

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). 1063 - 1070.

WebHamster (2009) Personal Space [image]. Available from https://commons.wikimedia.org/w/index.php?curid=6147809 [accessed 18th August 2017]

Wiederhold, B.K. Davis, R. Wiederhold, M.D. (1998) The Effects of Immersiveness On Physiology. In: Virtual Environments in Clinical Psychology and Neuroscience. Eds Riva G. Wiederhold, B.K. Molinari, B. IOS Press: Amsterdam.

Wiggins, J.S. (eds.) (1996) The five-factor model of personality: Theoretical perspectives. NY, USA: Guilford Press.

Wilcox, L.M., Allison, R.S., Elfassy, S. and Grelik, C. (2006). Personal space in virtual reality. *ACM Transactions on Applied Perception (TAP)*, *3*(4). 412-428.

Wired (2015) *THE HOBBIT: Battle of Five Armies VFX Crowd-Simulation On a Behemoth Scale | Design FX* [Video]. Available from https://www.youtube.com/watch?v=0NxSwskqwuY [Accessed 8th August 2017]

Witmer, B.G. and Singer, M.J. (1998). Measuring presence in virtual environments: A presence questionnaire. *Presence: Teleoperators and virtual environments*, 7(3). 225-240.

Zanlungo, F., Ikeda, T. and Kanda, T. (2011) Social force model with explicit collision prediction. *EPL journal* (Europhysics Letters), 93(6). 68005.

Zhou, S. et. al. (2010) 'Crowd modelling and simulation technologies'. *ACM Transactions on modelling and Computer Simulation,* 20(4).

10 Appendices

10.1 Appendix 1

Crowd Simulation in VR – Medical Screening Form

We operate this study according to the University of Lincoln School of Computer Science health and safety guidelines for Virtual Reality (VR) equipment. However, before you take part, it is important to determine whether you have any conditions which might impair your ability to use the VR equipment safely.

Please circle either 'yes' or 'no' to answer the following questions. If you need any help or wish to ask for any clarification, please ask:

Do you suffer from Epilepsy, or similar condition which may be triggered by flashing lights or visual stimulus?	YES	NO
Do you suffer from any significant uncorrected problems with your vision, such as tunnel vision? (if you wear glasses or contact lenses, answer 'no').	YES	NO
Are you pregnant? or likely to be pregnant?	YES	NO
Do you suffer from any conditions (e.g. related to mobility) which could cause you to be unduly injured by bumping into objects, or people, or by falling to the floor?	YES	NO
Do you suffer from Claustrophobia?	YES	NO
Do you suffer from any other condition which you think might affect your ability to use VR?	YES	NO

Please do ask if you would like to discuss anything relating these questions. This form will be destroyed at the end of the study.

10.2 Appendix 2

Crowd Simulation in VR – Participant Demographic Form

Please answer the questions on this form by circling or filling in where necessary.

Your gender	**MALE**	**FEMALE**	**OTHER**	**PREFER NOT TO SAY**
Your age				
Are you a University of Lincoln Student, University Staff, or neither?	**STUDENT**	**STAFF**	**BOTH**	**NEITHER**
Have you used a VR device before?	**YES**		**NO / NOT SURE**	
If you answered yes, which have you used?	**HTC VIVE** **PLAYSTATION VR** **OTHER:**		**OCCULUS RIFT** **MOBILE** **……………………………………**	
Do you play video games? If so which kind do you play most?				
On average how long do you spend playing video games a week (hrs)?				
Do you currently or have you in the past played first person video games?				

10.3 Appendix 3

Crowd Simulation in VR – PANAS 1

This scale consists of a number of words that describe different feelings and emotions. Read each item and then list the number from the scale below next to each word. Indicate to what extent you feel this way **right now**, that is, at **the present moment**.

1	2	3	4	5
Very Slightly or Not at All	A Little	Moderately	Quite a Bit	Extremely

_______________1.	Interested	_______________11.	Irritable
_______________2.	Distressed	_______________12.	Alert
_______________3.	Excited	_______________13.	Ashamed
_______________4.	Upset	_______________14.	Inspired
_______________5.	Strong	_______________15.	Nervous
_______________6.	Guilt	_______________16.	Determined
_______________7.	Scared	_______________17.	Attentive
_______________8.	Hostile	_______________18.	Jittery
_______________9.	Enthusiastic	_______________19.	Active
_______________10.	Proud	_______________20.	Afraid

10.4 Appendix 4

Crowd Simulation in VR – Participant Consent Form

Researcher: Jacob Greenwood
Researcher Email: JGreenwood@lincoln.ac.uk

Supervisor: Dr Patrick Dickinson

Please read the attached Participation Information Sheet, and then circle either 'yes' or 'no' to answer the following questions.

1	I have read and passed the medical screening form.	YES	NO
2	I have read and understood the Participation Information Sheet.	YES	NO
3	Any questions I have about the study at this point have been answered.	YES	NO
4	I understand that my participation in this study is voluntary.	YES	NO
5	I understand that I am free to withdraw at any point, including after I have completed the study today, without giving a reason.	YES	NO
6	I understand that my data will be treated confidentially, and held securely. Any publication resulting from this work will not include images or text which could identify me, without my express permission being sought.	YES	NO
7	I am 18 years of age, or older.	YES	NO

For participant to complete

Participant's Name	
Participant's Email	
Participant's Signature	
Date	

For Researcher to complete

Participant ID	

Note: This is the only sheet that contains personally identifying information. This sheet is the only link to the research data on the following pages, linked (one way only) by the ParticipantID. This sheet will be stored separately from the research data.

You will receive a copy of the information sheet and more information about the research at the end of the study. If you have any complaints or concerns about this research, you can direct these, in writing, to the Chair of the Research Ethics Committee by email at: ethics@lincoln.ac.uk

Participation Information Sheet

Project Summary: Investigating User Interactions in a Virtual Reality Crowd Simulation

You have been invited to take part in a research study which investigates human behaviour in a virtual reality crowd simulation. During the study, you will be asked to participate in a virtual reality world, and perform a number of simple tasks.

The study will typically take between 45 minutes and 1 hour for a single session.

During the study, we wish to collect the following data:

1. Some basic demographic data about you
2. Responses to a short questionnaire at a number of points in the study which measures your emotional state.
3. Video recordings of you while you are using the virtual reality equipment
4. Measures of your position and orientation in the virtual world, as you move around.
5. At the end of the study, we will ask you some questions about your experience. We will record your answers using an audio recording device.

This study is conducted in accordance with the University if Lincoln Ethical guidelines for participant studies, and has been approved by the College of Science Ethics Board.

Your Rights

- You may withdraw your participation at any time, without prejudice and without providing an explanation.
- If you wish to withdraw after the study has been completed today, you may ask that any data you have supplied, or that we have recorded, is removed. We undertake to destroy any data relating to you, should this circumstance arise.
- You may ask any questions that you have about the study or your participation, and we undertake to answer it as fully as possible.

Health and Safety

This study poses little risk to you as a participant. The study will be conducted according to the University of Lincoln School of Computer Science Health and Safety Guidelines which has included an appropriate risk assessment for virtual reality equipment.

Information Handling

Individual research data, including video and audio recordings, will only be viewed by members of the research team, and will be stored anonymously.

Any personally identifying information will be stored separately from the research data and can only be linked to research data through your unique participant ID (which is also held securely).

During the study you will be filmed from a visible stationary camera, this will only be recording whilst you are interacting with the virtual world. The video recorded will be kept confidential and secure. It will be transcribed using annotations by the researchers, for use in the study, and destroyed within 6 months after completion of the study.

If you are interviewed at the end of the study then your audio recording will be kept confidential and secure. It will be transcribed using annotations by the researchers, for use in the study, and destroyed within 6 months after completion of the study.

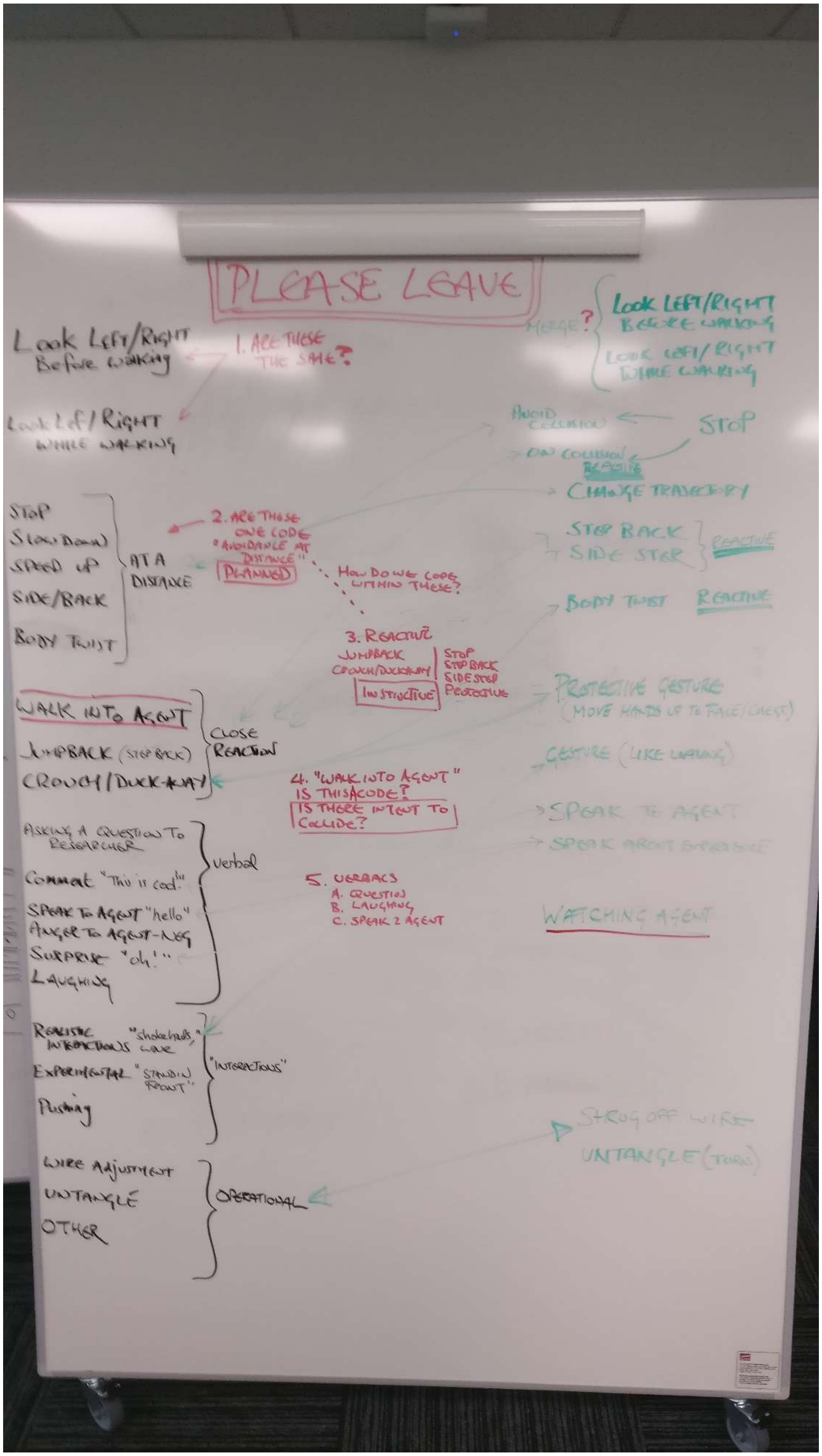
Images from video or quotes from audio recordings may be presented at conferences and in academic literature. However all such data will be anonymised, and if any video is used it will be re-acted so that you, the participant, are not identifiable.

The data gathered in this project will only be used for analysis relating to this project.

Further Information

Jacob Greenwood will be glad to answer any questions about the study at any time and can inform you about the results of the study once data collection is complete. You may contact him at **JGreenwood@lincoln.ac.uk**

10.5Appendix 5



10.6 Appendix 6

Category	No.	Code		Description
Look	1	Look left/right before walking	Lls	Clear look left and right before walking. Comparable to crossing a road
	2	Look left/right while walking	Llw	Looking left and right while walking
	3	Watching Agents	Lw	Participants watching a single agent for a couple seconds or more.
action at distance	4	Sudden Stop	Ast	Quick deceleration to stop while crossing
	5	Slow down	Asd	Gradual deceleration while crossing. Can include a stop if gradual enough
	6	Speed up	Asu	Speeding up to cross in front of agents
	7	Change Direction	Acd	Walking around agents
	8	Wait	Aw	Wait stationary while agents pass
close reaction	9	reaction stop	Rst	Impulsive stop
	10	step away	Rsa	Impulsive step away
	11	protective gesture	Rpg	Any gesture that could be protective, examples include: lifting hands up to face, slight crouching
	12	avoidance gesture	Rag	Moving hands out of way of agent without any clear body movement
	13	Body twist	Rbt	Twisting body to make room for agent to pass.
	14	Minor Reaction	Rm	Impulsive reaction that is too slight for other codes.
	15	Reaction Speed Up	Rsu	Suddenly speeding up to avoid agents, similar to Rsa
	16	Reaction Slow Down	Rsd	Suddenly slowing down without stopping.
physical interaction	17	Touch	It	Touching an agent either experimental or interaction
	18	Push	Ip	Trying to push agents
verbal	19	speak to agent	V2a	“hello” “move out of the way”
	20	about agent	V2r	“there’s lots of them”
	21	spontaneous utterance	Vsu	Usually linked to reaction and includes any shocked noises that aren’t words
	22	laughing	Vl	Nervous or positive laughter
Operational	23	Wire adjustment	Owa	Minor wire adjustments such as ducking shoulder under wire.
	24	Wire management	Ou	Clear stop and adjustment, such a turning whole body. Anything else that happens during this time is not coded.
	25	ask researcher	Oar	Any non-rhetorical question. “Can I start?” “Can I reach that?”
Other	26	special note	Spc	Special note for researcher to look back on later

10.7 Appendix 7

Code		Low	Medium	High
Look left/right before walking	lls	96	86	99
Look left/right while walking	llw	116	117	94
Watching Agents	lw	43	39	12
Sudden Stop	ast	25	32	61
Slow down	asd	33	56	53
Speed up	asu	6	4	4
Change Direction	acd	41	68	88
Wait	aw	55	44	39
reaction stop	rst	19	24	53
step away	rsa	11	10	24
protective gesture	rpg	1	2	9
avoidance gesture	rag	3	11	20
Body twist	rbt	3	5	8
Minor Reaction	rm	20	16	41
Reaction Speed Up	rsu	1	0	0
Reaction Slow Down	rsd	1	0	3
Touch	it	1	0	0
Push	ip	1	3	4
speak to agent	v2a	1	2	4
about agent	v2r	1	5	6
spontaneous utterance	vsu	7	5	17
laughing	vl	2	2	2
Wire adjustment	owa	2	3	8
Wire management	ou	6	2	8
ask researcher	oar	4	1	3
special note	spc	1	2	3

10.8 Appendix 8

Positive PANAS Scores					Negative PANAS Scores				
Baseline	Training	Low	Medium	High	Baseline	Training	Low	Medium	High
42	45	37	41	24	14	13	21	15	36
37	45	43	44	37	10	10	12	10	12
29	33	33	33	35	10	10	10	10	10
34	34	27	31	28	13	10	10	13	13
28	37	30	34	34	10	10	10	11	13
43	47	46	47	46	11	15	15	16	20
27	32	30	26	28	11	10	13	11	13
32.5	39	39	31	36	10	14	18	11	10
38	42	38	39	30	20	21	19	28	32
28	33	26	33	31	17	10	14	19	21
24	23	25	21	28	19	12	10	10	18
32	30	29	28	28	13	11	11	12	14
35	40	36	34	33	11	10	11	14	13
22	18	16	16	11	11	10	13	17	17
31	36	35	32	33	10	10	11	13	11
32	41	34	37	33	10	10	10	15	21
20	29	33	29	28	10	10	10	13	15
33	36	36	37	35	12	10	10	10	10
41	46	45	45	44	13	13	13	14	15
26	26	21	21	21	10	10	10	10	10
42	46	46	46	44	15	10	10	10	12
40	40	43	45	41	16	14	16	14	11
34	46	48	47	43	16	10	11	10	15
21	25	20	19	22	11	10	15	15	18
38	44	48	47	47	11	11	11	15	14