2013

The Turning, Stretching and Boxing Technique: a Direction Worth Looking Towards

Mark Dunne
Technological University Dublin, mark.dunne@tudublin.ie

Recommended Citation

Dunne, M. (2013). *The turning, stretching and boxing technique: a direction worth looking towards* Doctoral Dissertation. Technological University Dublin. doi:10.21427/D71P4X

The *Turning*, *Stretching* and *Boxing* Technique: A Direction Worth Looking Towards

by

Mark Dunne

Supervisors: Dr. Brian Mac Namee
Dr. John Kelleher

School of Computing

Dublin Institute of Technology

A thesis submitted for the degree of

Doctor of Philosophy

April, 2013

This is dedicated to my parents, James & Margaret.

Thank you both for your unwavering support throughout the years.

Declaration

I certify that this thesis which I now submit for examination for the award of Doctor of Philosophy, is entirely my own work and has not been taken from the work of others, save and to the extent that such work has been cited and acknowledged within the text of my work.

This thesis was prepared according to the regulations for postgraduate study by research of the Dublin Institute of Technology and has not been submitted in whole or in part for an award in any other Institute or University.

The work reported on in this thesis conforms to the principles and requirements of the institutes's guidelines for ethics in research.

The Institute has permission to keep, to lend or to copy this thesis in whole or in part, on condition that any such use of the material of the thesis be duly acknowledged.

Signature__________________________ Date__________________________

Acknowledgements

The Supervisors: I owe a huge depth of gratitude to Dr Brian Mac Namee and Dr John Kelleher for all their support, advice and perseverance throughout my PhD journey.

The Participants: I would like to thank all the participants for the generosity of their time, their patience and most importantly the excellent data they provided.

The Editor: I would like to thank Emma O'Dwyer for her diligence at editing this thesis, your time and patience was greatly appreciated.

My Girlfriend: Niamh Parkinson, I would especially like to thank you for all the support and encouragement you gave to me over the last few years of my PhD journey, without it I would have most likely not made it this far.

Abstract

3D avatar user interfaces (UI) are now used for many applications, a growing area for their use is serving location sensitive information to users as they need it while visiting or touring a building. Users communicate directly with an avatar rendered to a display in order to ask a question, get directions or partake in a guided tour and as a result of this kind of interaction with avatar UI, they have become a familiar part of modern human-computer interaction (HCI). However, if the viewer is not in the sweet spot (defined by Raskar *et al.* (1999) as a stationary viewing position at the optimal 90° angle to a 2D display) of the 2D display, the 3D illusion of the avatar deteriorates, which becomes evident as the user's ability to interpret the avatar's gaze direction towards points of interests (PoI) in the user's real-world surroundings deteriorates also.

This thesis combats the above problem by allowing the user to view the 3D avatar UI from outside the sweet spot, without any deterioration in the 3D illusion. The user does not lose their ability to interpret the avatar's gaze direction and thus, the user experiences no loss in the perceived *corporeal* presence (Holz *et al.*, 2011) for the avatar. This is facilitated by a three pronged graphical process called the *Turning*, *Stretching* and *Boxing* (TSB) technique, which maintains the avatar's

3D illusion regardless of the user's viewing angle and is achieved by using head-tracking data from the user captured by a Microsoft *Kinect*.

The TSB technique is a contribution of this thesis because of how it is used with an avatar UI, where the user is free to move outside of the sweet spot without losing the 3D illusion of the rendered avatar. Then each consecutive empirical study evaluates the claims of the TSB Technique are also contributions of this thesis, those claims are as follows: (1) *increase interpretability of the avatar's gaze direction* and (2) *increase perception of corporeal presence for the avatar*. The last of the empirical studies evaluates the use of 3D display technology in conjunction with the TSB technique.

The results of *Study 1* and *Study 2* indicate that there is a significant increase in the participants' abilities to interpret the avatar's gaze direction when the TSB technique is switched on. The survey from *Study 1* shows a significant increase in the perceived corporeal presence of the avatar when the TSB technique is switched on. The results from *Study 3* indicate that there is no significant benefit for participants' when interpreting the avatar's gaze direction with 3D display technology turned on or off when the TSB technique is switched on.

Contents

List of Tables

List of Figures

Chapter 1

Introduction

Avatar user interfaces (UI) are a means to visually represent an *intelligent virtual agent* (IVA) on a medium such as a 2D display, to a user who can then communicate with the avatar in a natural way. An avatar UI is useful in many scenarios such as its use in large public spaces like museums, where they can add value to a user's experience of being in a new environment by delivering location based information.

A good example of this is *MIKI* (McCauley & D'Mello, 2006), an assistive agent that can appear in a kiosk within the lobby of a public building. *MIKI* can answer a user's questions or direct them to wherever they need to go within the building by displaying a map on the large 50 inch 2D display and can give clear verbal directions. *MIKI* has a life-like human avatar which appears in the bottom left of the 2D display. *MACK* (Cassell *et al.*, 2002) is a similar kiosk-style avatar UI to *MIKI* which also provides visitors to a public building with relevant information but instead of a human-like avatar, *MACK* has an animated robot as the avatar. *MIKI* (see Figure

1.1) and *MACK* (see Figure 1.2) both provide mediated experiences (Riva *et al.*, 2003, pp.3-16) (e.g. watching a film in the cinema is a mediated experience) due to the fact that the 2D displays on which they render their avatars are clearly evident in their respective kiosks within their users' immediate vicinity.

Both *MIKI* and *MACK* are not capable of delivering accurate referencing gestures (e.g. pointing) (Kita, 2003) or referencing gaze behaviours (Lobmaier *et al.*, 2006) from their 2D display in order to reference anything in the user's immediate vicinity. This type of referencing is an integral part of human-to-human communication (Garau *et al.*, 2001; Kita, 2003; Tomasello, 2008) and as such, would be helpful to a user while the avatar is giving them information about their immediate environment. The mediated nature of these kiosk-styled avatar UIs can affect the level of immediacy and intimacy the user experiences when interacting with the avatar. Therefore, the user becomes aware of the medium (i.e. 2D display) on which the mediated interaction is occurring and their sense of the avatar's *social presence* (Lombard & Ditton, 1997) (or *presence*) in their real-world surroundings is greatly diminished. This thesis is particularly focused on the concept of *corporeal* presence (Holz *et al.*, 2011) as it directly relates to the user's perception of the avatar's body (i.e. 3D representation of the avatar) within the user's immediate vicinity and is less concerned with the social aspects of presence.

Lombard *et al.* (2000, p.1) define presence as '*the perceptual illusion of non-mediation*', where the term *perceptual* indicates the continuous (i.e. real-time) responses of the user (i.e. *sensory*, *cognitive* and *affective* processing systems) to objects and entities in their real-world surroundings. The *illusion of non-mediation* occurs when the user fails to perceive or acknowledge the medium (i.e. the 2D

Figure 1.1: The MIKI kiosk set-up (McCauley & D'Mello, 2006).

Figure 1.2: A user interacting with MACK (Cassell *et al.*, 2002), the scene behind the 2D display is rendered to the screen to make it look as if *MACK* is standing within the user's environment.

display on which the avatar is displayed) in their real-world surroundings and they respond as they would if the medium was not there. The consequence of a user being able to ignore the medium benefits their interaction with the avatar by allowing for immediacy and intimacy to occur, both of which play an important role in face-to-face interactions between social agents leading to *connectedness* (Rettie, 2003).

This thesis will present a graphical approach for displaying an avatar UI on a standard 2D display, which will be continuously updated to match the user's perspective in order to achieve a greater sense of corporeal presence for the avatar while allowing the user to freely move outside of the sweet spot of the 2D display. This graphical approach should enhance the perceived realism of the avatar as it is continuously rendered to reflect the user's perspective, in turn helping to create '*the perceptual illusion of non-mediation*' (Lombard *et al.*, 2000) and establishing a greater sense of the avatar's corporeal presence in the user's vicinity. This has two benefits, first, the user is not restricted to viewing the avatar from the sweet spot as seen in the *MIKI* (McCauley & D'Mello, 2006) and *MACK* (Cassell *et al.*, 2002) kiosk-style avatar UIs. Second, the user can accurately interpret the avatar's gaze direction as it looks out of the 2D display into the user's real-world surroundings. Overall, the use of the above graphical approach should contribute to user's increased perception of the avatar's presence, and more specifically the perceived corporeal presence of the avatar.

Next there will be an introduction to the several views of presence as outlined by Holz *et al.* (2011); Lombard & Ditton (1997); Slater (2009) and an illustration as to how they relate to each other in Figure 1.3. Holz *et al.* (2011) defines corporeal presence as having two contributing factors which are as follows:

1. *Representation*: Describes the perceived representation of the avatar in the real-world, in the case of an avatar UI this representation manifests itself as a virtual body rendered to a 2D display in the user's vicinity.

2. *Geometric Correspondence*: The perceived relationship between the representation of an avatar UI as it resides in a virtual sub-space (i.e. the projection of the avatar on the 2D display) and how this representation corresponds

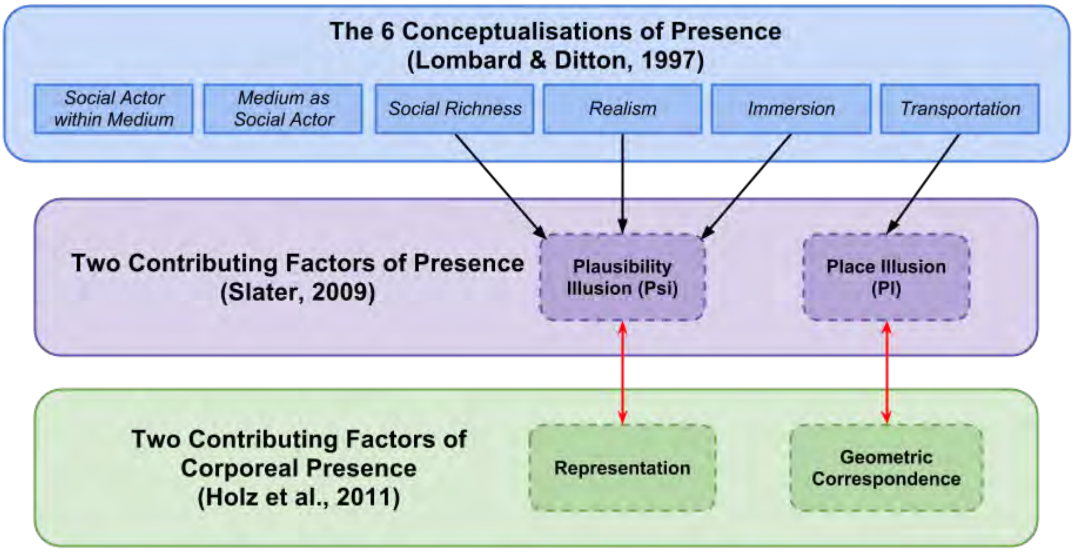


Figure 1.3: This diagram illustrates the link between six *concepts of presence* put forth by Lombard & Ditton (1997) and *corporeal presence*, a more refined concept of presence specifically relating to Mixed Reality Agents (MiRA) outlined by Holz *et al.* (2011). Corporeal presence has two main factors, *representation* and *geometric correspondence*. These two contributing factors of corporeal presence can be directly related to the two contributing factors of presence, *plausibility illusion* (Psi) and *place illusion* (Pl), discussed by Slater (2009). Psi and Representation directly correspond to each other, as do Pl and *geometric correspondence* as indicated by the red arrows. The black arrows show the relationship between the six conceptualisations of presence outlined by Lombard & Ditton (1997) and the two contributing factors of presence according to Slater (2009). Suffice to say, the same conceptualisations of presence that are related to Psi and Pl also relates to the two contributing factors of corporeal presence by Holz *et al.* (2011), *representation* and *geometric correspondence*.

geometrically to the physical sub-space (i.e. the real-world containing a 2D display with an avatar rendered to it).

Displaying the correct visual representation of the avatar for the user's perspective (i.e. the user's viewing angle) not only adds credence to the 3D illusion of the avatar but can also lead to the user's increased accuracy at interpreting the avatar's gaze direction. This is because the avatar's gaze direction correctly corresponds geometrically to the real-world as the avatar looks out of its virtual sub-space from the user's perspective. This correct correspondence grounds the avatar into the user's real-world surroundings and hence, increasing the perceived corporeal presence of the avatar. Furthermore, the two contributing factors of corporeal presence can be directly linked to two contributing factors of presence outlined by Slater (2009): *Plausibility Illusion* (Psi) and *Place Illusion* (Pl). Slater (2009) describes Psi as the user's belief that what is happening is actually happening and Pl as the belief the user has of being transported to a virtual world in relation to Virtual Reality (VR). However, in terms of a MiRA (Holz *et al.*, 2011), Pl correlates to the level of accurate geometric correspondence an avatar UI can achieve with the user's real-world surroundings and therefore, it also could be interpreted as a level of the user's belief that the avatar is co-habiting their real-world surroundings.

Lombard & Ditton (1997) have a conceptualisation of presence for the user's belief of co-habitation called *transportation*, which relates to how the user perceives themselves in the context of another social agent: (1) '*you are there*', (2) '*it is here*' and (3) '*we are together*'. The scope of this research is concerned with the second description of transportation, i.e. '*it is here*' (the avatar is the '*it*'). Psi has a direct bearing on three of the six conceptualisations of presence as outlined by Lombard & Ditton (1997), *Social Richness* (see Section 2.1.3), *Realism* (see Section 2.1.1) and *Immersion* (see Section 2.1.2). As there is a direct relationship between Psi and representation as described by Holz *et al.* (2011), the three conceptualisations of presence listed above subsequently can be related to representation which is illustrated in Figure 1.3.

As the illusion of the avatar's body becomes more established in the user's physical reality, the user will increasingly perceive that the avatar really exists in their real-world surroundings, i.e. '*it is here*'. This is beneficial for an avatar UI as it creates social richness, immersion and realism for the user while interacting with the avatar. All three are important for an avatar UI as such interfaces tend to require a user to socially engage with an avatar. Hence, any increase in perceived corporeal presence may lead to a subsequent increase in the more general idea of perceived presence of the avatar as a social entity (i.e. the avatar's ability to encourage engagement/communication from a user).

1.1 Problem Domain

The standard approach to displaying a 3D avatar UI is to render it onto 2D displays (e.g. LCD panel, projector, mobile device) as seen in other research such as *MACK* (Cassell *et al.*, 2002) and *MIKI* (McCauley & D'Mello, 2006), or project the avatar UI directly onto a wall like the *Virtual Room Inhabitant* (Kruppa *et al.*, 2005). However, most 2D displays require the user to remain in a stationary position at the optimal 90° viewing angle both horizontally and vertically perpendicular to the display. This optimal viewing angle is commonly referred to as the *sweet spot*, as it is the point where images are created for the ideal single viewer location (i.e. the static user as described by Raskar *et al.* (1999)) as seen in Figure 1.4 which outlines the sweet spot area in front of the 2D display.

Any viewing angle more acute than the sweet spot viewing angle can cause deterioration in the effectiveness of the rendered 3D graphics and thus the 3D illusion it creates. Simply put, the image is distorted from the user's perspective, a phenomenon known as *lateral foreshortening*. Lateral foreshortening can become problematic for maintaining 3D illusions at viewing angles greater than 75° from the sweet spot's optimal viewing angle. This occurs when the user is viewing an image from an angle that exceeds the limits of the image's visual field. However, the 3D illusion starts to a gradually deteriorate and this deterioration become more apparent as the user's viewing angle strays further from the sweet spot. This in turn effects the perceived corporeal presence of the avatar. In the empirical studies carried out in this thesis (see Chapters 4, 5 and 6) it becomes evident that this gradual deterioration in the 3D illusion inhibits the user's ability to use their own gaze perception (see Section 2.2) to interpret the avatar's gaze direction. This ability is needed to help the user accurately determine where the avatar is directing its gaze out into the user's real-world surroundings. The avatar no longer correctly corresponds geometrically to the user's real-world surroundings from that user's perspective. This means the user's ability to accurately interpret the avatar's gaze direction, which adds realism to the interaction by creating corporeal presence, is lost and a long with it '*the illusion of non-mediation*' (Lombard & Ditton, 1997) is also lost. Subsequently, the user's interpretation of other gaze behaviours by the avatar can be effected by the loss in corporeal presence and as gaze behaviours are important indicators of the willingness of one social agent to engage in social interactions with another (Peters *et al.*, 2005) this can have detrimental effect.

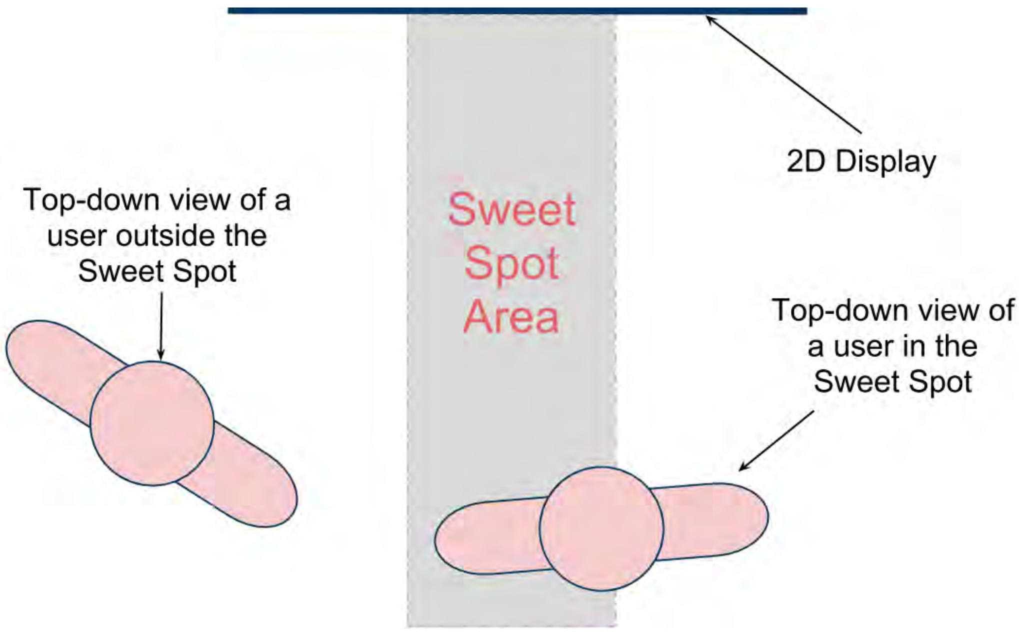


Figure 1.4: This diagram depicts a top-down view of the *Sweet Spot*, the optimal viewing position at a 90° angle to the centre of the 2D display. There are two users in the diagram, one illustrating a viewing position inside of the sweet spot (at an almost 90° viewing to the centre of the 2D display) and another illustrating a viewing position outside of the *sweet spot*. The further from the *sweet spot* a user views the rendered 3D image of the avatar on the 2d display, the more the 3D illusion of the avatar diminishes. This is caused by the increasingly obvious distortion of the 3D illusion due to lateral foreshortening as the user strays further from the *sweet spot*.

Lateral foreshortening is the main cause for the decrease in the user's perceived corporeal presence of the avatar. As the 3D illusion diminishes the medium (i.e. the 2D display) on which the avatar is displayed becomes more apparent. This occurs when the medium delivering the mediated interaction can no longer be ignored and hence, becomes a barrier to a natural communication style between the user and the avatar. The user loses their sense of immersion as they become more aware that the avatar is only a projection and is not actually there with them, co-existing in their real-world environment. Therefore, preventing lateral foreshortening and the subsequent decrease in corporeal presence can in turn decrease the user's awareness of the medium on which the avatar is displayed. This is similar to a notion commonly instilled through use of virtual reality where a user becomes so immersed in the virtual world that they feel the sense of *being there* (Sas & O'Hare, 2003; Witmer & Singer, 1998). Any decrease in corporeal presence will correlate with any decreases in Pl and Psi during a user's experience. Hence, addressing the issue of lateral foreshortening by attempting to create and maintain corporeal presence could potentially lead to an increased sense of social richness, realism and immersion during interactions with an avatar (Lombard & Ditton, 1997).

There is a need to address the issue that arises when displaying an avatar UI on a 2D display. The interaction between the user and the avatar can suffer due to the restriction put on the user to stand in the sweet spot or it can deteriorate further because as the user moves away from the sweet spot the 3D illusion diminishes. Ultimately this leads to a decrease in the user's ability to accurately interpret the avatar's gaze direction and subsequently, contributes to a decrease in the user's perceived corporeal presence of the avatar. These decreases have the potential to negatively effect the natural communication style that would otherwise be established and sustained.

1.2 Contributions of this Thesis

In order to tackle the problem domain as outlined above in Section 1.1, the development of a graphical framework called the *Turning*, *Stretching* and *Boxing* (TSB) technique (see Chapter 3) that uses the Microsoft Kinect to track a user's head position was carried out. The TSB technique uses head-tracking data captured from the Kinect to determine the eye position of the tracked user and then uses this data to render the 3D avatar onto a 2D display so that the 3D illusion of the avatar is maintained regardless of the user's viewing angle (i.e. the user's perspective of the 2D display).

The Microsoft Kinect uses an infra-red laser projector and camera to track up to 20 joints (i.e. features) per user (i.e. up to 2 users) in real-time and has a Field of View (FoV) of only 57°. The accuracy of the sensor while interpreting depth can vary as much as a few millimetres up to 4 cm when the user being tracked is at the maximum range (Obdrzalek *et al.*, 2012) (i.e. over 4 metres). The Kinect can calibrate itself for multiple lighting conditions and can auto adjust its vertical viewing angle using an internal motor to allow for differences in the height of tracked users (i.e. the Kinect will adjust to ensure it is tracking the user fully, as long as the user stays within its FoV and range).

Individually the *turning*, *stretching* and *boxing* processes are not novel graphical techniques and have been used to some extent before in other avatar UI research (Agrawala *et al.*, 1997; Kipp & Gebhard, 2008). The "*Responsive Workbench*" by Agrawala *et al.* (1997) (see Figure 1.5 and Figure 1.6) allows up to two users to view the surface of the Workbench perspectively correct for each of their eyes (i.e. renders 4 images, one for each eye) by stretching the 3D scene to account for lateral foreshortening of viewing the 2D display from outside of the sweet spot and this leads to each user experiencing a stereoscopic image. Kulik *et al.* (2011) further expand on the research by Agrawala *et al.* (1997) by allows 6 users to view a 3D scene simultaneously, with each user getting a perspectively correct image for each of their eyes. The TSB technique is limited to one tracked user and the imaged rendered to the 2D display is not stereoscopic. For the TSB technique to work with multiple users, the users would have to wear 3D glasses (i.e. passive or active 3D glasses) just like the users in the research by both Agrawala *et al.* (1997) and Kulik *et al.* (2011).

However, the TSB technique is still a contribution of this thesis because how it is evaluated for its use with an avatar UI, where the tracked user is able to freely move outside of the sweet spot for the 2D display without losing the 3D illusion of the rendered avatar UI. The two claims made for the benefits of using the TSB technique in conjunction with an avatar UI on a 2D display where the user is not restricted to the sweet spot, are as follows:

Figure 1.5: In this image two users can be seen to simultaneously interact with a 3D cube while viewing a shared virtual environment known as the the "*Responsive Workbench*". The calibration of the Workbench ensured that when the two users point to the same feature (i.e. the same corner) on the virtual cube, their fingers touch. However, for the purposes of this image Agrawala *et al.* (1997) manipulated the image on the Workbench so it is rendered from the point of view of the camera and subsequently the viewer of this image. This gives the impression of what the actual users saw. Hancock *et al.* (2009) also discuss this method of rendering a 3D image on a 2D surface to match the users' perspectives.

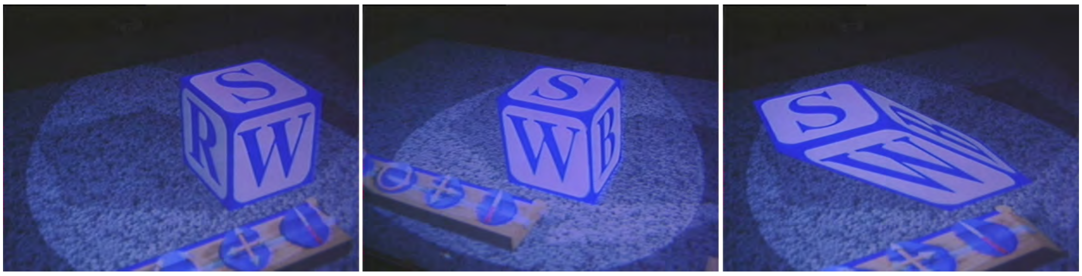


Figure 1.6: The image to the left is of the rendered scene on the Workbench for a user's perspective who is standing on the left side of the table. The image in the centre is from a user's perspective who is standing to the right side. The image on the right is the view that a user standing on the left side of the table would see if the 3D scene was generated for the perspective of a user standing to the right of the Workbench and it suffers from lateral foreshortening. Even though the image on the Workbench is the same in both the centre and right images, the cube appears sheared to the non-tracked user.

1. **Increase interpretability of the avatar's gaze direction**: The TSB technique increases a user's accuracy at interpreting the avatar's gaze direction towards *points of interest* (PoI) in the user's real-world surroundings from the user's perspective. This increase is most noticeable when the user is interpreting the avatar's gaze direction from outside of the sweet spot of the 2D display.

2. **Increase perception of corporeal presence for the avatar**: The rendering of the 3D avatar is always correct for the tracked user's perspective, ensuring the visual fidelity of the rendered avatar as the user moves in front of the 2D display. In turn this visual fidelity helps to establish the avatar's representation within the user's environment. In addition, the visual fidelity of the rendered avatar means the avatar correctly corresponds geometrically to the user's environment from the user's perspective. Leading to the user's higher level of perceived corporeal presence for the avatar.

The above two claims were evaluated in a series of three empirical studies (see Chapter 4, 5 and 6) and these evaluations form the basis of three additional contributions for this thesis, they are as follows:

1. **Evaluation to test the effect of the TSB technique on a user's ability to interpret the avatar's gaze direction (see Chapters 4 and 5)**: The TSB technique increases the ability of a user's own *gaze perception* (Monk & Gale, 2002) to accurately interpret the avatar's gaze direction as it looks out of the 2D display it is rendered on towards PoI in the user's real-world surroundings. Furthermore, from the perspective of the user this facilitates a higher level of perceived interaction between the avatar and the real-world where the 2D display with the rendered avatar is placed. Thus, establishing '*the perceptual illusion of non-mediation*' (Lombard *et al.*, 2000, p.1), where the 2D display is perceived to be more like a real-world window and behaves as such. The results of the empirical *Study 1* (see Chapter 4), *Study 2* (see Chapter 5) and *Study 3* (see Chapter 6) show that the recorded levels of accuracy by participants at interpreting the avatar's gaze direction were significantly higher during gaze perception trials when the TSB technique is switched on. These results help establish the first claim of the TSB technique and have been published in Dunne *et al.* (2012b).

2. **Evaluation to test the effect of the TSB technique on a user's perception of the avatar's corporeal presence (Chapter 4)**: This is due to the increased level of Psi experienced by the user as the sustained 3D illusion of the avatar further increase the feeling of the avatar being present in

the vicinity of the user. In turn, this helps the user to ignore the medium (i.e. a 2D display) that mediates the interaction with the avatar, potentially heightening the levels of realism, immersion and social richness the user experiences. Previous research by Lombard & Ditton (1997) indicates a link between increased interaction (i.e. the avatar's ability to accurately use gaze direction to look at PoI in the user's real-world surroundings) and perceived presence. As interaction between the avatar and the user's real-world surroundings becomes more realistic the user experiences increases in their perceived corporeal presence for the avatar. A survey was used in *Study 1* (see Section 4.3) to evaluate the second claim of the TSB technique. The results of the survey indicate that the increase in the avatar's realistic representation and the increased interaction due to the avatar corresponding geometrically correct to the participant's real-world surroundings does have an impact on the user's level of perceived corporeal presence for the avatar. The results of this study help establish the second claim of the TSB technique and have been published in Dunne *et al.* (2012b).

3. **Evaluation to test if 3D display technology has any bearing on a user's ability to interpret the avatar's gaze direction with and without the TSB technique being switched on (Chapter 6)**: When a user views an avatar UI with 3D display technology turned on (i.e. see Anaglyph 3D, Section 6.1), from outside of the sweet spot, a graphical technique such as the TSB technique is still required in order to allow accurate interpretation of the avatar's gaze direction. Also, there is no impact on the user's ability

to accurately interpret the avatar's gaze direction towards a PoI in the user's real-world surroundings with 3D display technology alone. This is further evidence that use of 3D display technology does not eliminate the need for a graphical framework such as the TSB technique. Similar to a standard 2D display, a user viewing an avatar UI with 3D display technology turned on, is still required to remain in the sweet spot as it is the optimal viewing position where the 3D illusion of the avatar is at its strongest. An empirical study (see Chapter 6) was carried out to test if there was any impact on a participant's ability to accurately interpret the avatar's gaze direction when 3D display technology was used with and without the TSB technique being switched on. The findings of this empirical study showed that 3D technology did not impact a participant's accuracy at interpreting the avatar's gaze direction and these findings were published in Dunne *et al.* (2012a).

1.3 Overview

The remainder of this thesis is structured as follows: Chapter 2 will discuss the background for the research reported in this thesis by outlining the important areas such as corporeal presence, gaze perception and the Mona Lisa Effect. Corporeal presence derives from the 3D illusion of the rendered avatar and how it is manifested in the user's cognitive processes. Leading to higher accuracy at interpreting the avatar's gaze direction, creating shared references in the user's real-world surroundings and contributing to the user's overall belief that the avatar is actually cohabiting the user's real-world surroundings. The increased interaction between avatar and the user's real-world surroundings results in higher perceived presence, particularly for four of six conceptualisations of presence for the avatar as outlined by Lombard & Ditton (1997) (*Social Richness*, *Realism*, *Immersion* and *Transportation*). This is followed by a thorough discussion on how it relates all the concepts of presence are related (see Figure 1.3), which is important as the premise of this thesis is how to create and maintain a high level of corporeal presence for an avatar UI. Next a brief analysis of corporeal presence (Holz *et al.*, 2011) takes place with a discussion on the two contributing factors for presence (i.e. Psi and Pl) as outlined by Slater (2009). The remainder of Chapter 2 outlines the technical background for the TSB technique.

Chapter 3 outlines the three contributing graphical processes (i.e. *turning*, *stretching* and *boxing*) of the TSB technique, which all use head-tracking data from the Microsoft Kinect to render the avatar UI to a 2D display in the correct perspective to match the tracked user's viewing angle. This in turn increases the user's ability to interpret the avatar's gaze direction towards PoI in the user's real-world surroundings, helping to establish joint attention with shared references. Subsequently, the TSB technique allows for a high level of interaction between the avatar and the user's real-world surroundings, which results in an increased level of perceived corporeal presence for the avatar and this increase can be seen in the results of the survey carried out in *Study 1* (see Chapter 4).

Chapter 4 details the first of three empirical studies, *Study 1*, and evaluates both claims of the TSB technique: (1) increased interpretability of the avatar's gaze direction for the user; (2) the increase in perceived corporeal presence for the avatar felt by the user. In *Study 1* each participant had to guess which floor marker the avatar is directing its gaze towards at any one time. There are seven floor markers and each participant started each of the trials by standing on a floor marker. Next the avatar directs its gaze towards one of the remaining floor markers. This was done forty-two times ensuring that each participant guessed for all the possible combinations of moves between all seven of the floor markers for each of the two experimental conditions, a control condition and a TSB condition. After analysing the results from all participants in *Study 1* there is strong evidence to indicate that an increase in the participants' abilities to interpret the avatar's gaze direction does occur with the TSB technique switched on and supports the first claim of the TSB technique. Similarly, when the survey results from all participants in *Study 1* were statistically analysed, they indicated an increase in perceived corporeal presence for the avatar and support the second claim of the TSB technique.

Chapter 5 details *Study 2*, with a more elaborate experimental set-up than *Study 1* and provides a more rigorous evaluation of the first claim of the TSB technique (i.e. the increased ability of a user to interpret the avatar's gaze direction). In *Study 2* the participants were required to interpret the avatar's gaze direction in two ways: (1) the participants had to guess under which of the three upturned buckets a prize (i.e. sweets/candy) was hidden and (2) the participants had to guess what letter the avatar was gazing towards on the surrounding walls of the laboratory. The analyses of the data gathered from all the participants during each trial revealed that participants achieved a higher accuracy rating when the TSB technique was switched on. This result further backs up the first claim of the TSB technique and mirrors the results previously seen in *Study 1*.

Chapter 6 details the final empirical study in this thesis, *Study 3*, which evaluates whether or not the addition of 3D display technology has any bearing on a participant's ability to accurately interpret the avatar's gaze direction towards PoI in the participant's real-world surroundings. The experimental set-up for *Study 3* is identical to that of *Study 1*, however, with the addition of two other experimental conditions: (1) a second control condition with 3D technology switched on and (2) a second TSB condition with 3D technology switched on. Anaglyph 3D technology was used for the experimental set-up due to its ease of adoption and the availability of the low cost anaglyph 3D glasses. After analysing the results from all the participants in *Study 3*, it is clear that the use of 3D technology has no bearing on a participant's ability to interpret the avatar's gaze direction. This is true when the results were compared for both of the TSB conditions with and without 3D technology being turned on. The same outcome is seen between the comparison of the two control conditions with and without 3D technology being turned on. The results from this empirical study further support the first claim of the TSB technique and once again mirror the results seen in *Study 1*.

Finally, Chapter 7 begins with a discussion on our approach for evaluating the claims of TSB technique with some additional findings and observations outlined before moving on to future work. This chapter is concluded and the thesis is brought to a close with some final thoughts.

1.4 List of Publications

This thesis is supported by the following publications:

[**Dunne *et al.* (2012b)**] Dunne, M., Mac Namee, B. & Kelleher, J.: The turning, stretching and boxing technique: A step in the right direction. In Y. Nakano, M. Neff, A. Paiva & M. Walker, eds., *Intelligent Virtual Agents*, vol. 7502 of *Lecture Notes in Computer Science*, 363–369, Springer Berlin Heidelberg.

[**Dunne *et al.* (2012a)**] Dunne, M., Mac Namee, B. & Kelleher, J.: Stereoscopic avatar interfaces: A study to determine what effect, if any, 3d technology has at increasing the interpretability of an avatar's gaze into the real-world. In *Multimodal Analyses enabling Artificial Agents in Human-Machine Interaction (MA3 2012)*, Santa Cruz, Ca, USA.

[**Dunne *et al.* (2010b)**] Dunne, M., Mac Namee, B. & Kelleher, J.: TSB Technique: Increasing a User's Sense of Immersion with Intelligent Virtual Agents. *The 21st National Conference on Artificial Intelligence and Cognitive Science Student Symposium (AICS 2010).*

[**Dunne *et al.* (2010a)**] Dunne, M., Mac Namee, B. & Kelleher, J.: Scalable Multimodal Avatar Interface for Multi-user Environments. *The International Conference on Computer Animation and Social Agents 2010 (CASA 2010).*

[**Dunne *et al.* (2009)**] Dunne, M., Mac Namee, B. & Kelleher, J.: Intelligent Virtual Agent: Creating a Multi-Modal 3D Avatar Interface. In *Proceedings of the 9th Annual Information Technology & Telecommunication Conference (IT&T 2009).*

[**Mac Namee & Dunne (2009)**] Mac Namee, B. & Dunne, M.: Widening the Evaluation Net. In *Intelligent Virtual Agents (IVA 2009)*, 525–526, Springer Berlin Heidelberg.

As a summary, the contributions of this work, the corresponding chapters of this thesis and the publications are shown in Table 1.1.

Table 1.1: Contributions, corresponding chapters and publications.

Contribution	**Chapter(s)**	**Publication**
TSB Technique with the two claims: (1) increase interpretability of the avatar's gaze direction and (2) increase perception of corporeal presence for the avatar.	Chapter 3	Dunne *et al.* (2010b)
Evaluation to test the effect of the TSB technique on a user's perception of the avatar's corporeal presence	Chapter 4	Dunne *et al.* (2012b)
Evaluation to test the effect of the TSB technique on a user's ability to interpret the avatar's gaze direction	Chapter 4 and 5	Dunne *et al.* (2012b)
Evaluation to test if 3D display technology has any bearing on a user's ability to interpret the avatar's gaze direction with and without the TSB technique being switched on	Chapter 6	Dunne *et al.* (2012a)

Chapter 2

Background Literature

To briefly recap the contributions of this thesis (see Section 1.2), there is the use of the TSB technique with an avatar UI rendered on a 2D display, where the user is free to move outside of the sweet spot without losing the 3D illusion of the rendered avatar. Subsequently, increasing the perceived level of corporeal presence for the avatar in the user's real-world surroundings. Then each consecutive empirical study evaluates the claims of the TSB Technique and are also contributions, those claims are as follows: (1) *increase interpretability of the avatar's gaze direction* and (2) *increase perception of corporeal presence for the avatar*. The last of the empirical studies evaluates the use of 3D display technology in conjunction with the TSB technique. All three studies measure the participants' accuracy at interpreting the avatar's gaze direction.

In this chapter the relevant background literature is highlighted, starting with the six conceptualisations of presence by Lombard & Ditton (1997). Three of the six conceptualisations (i.e. *realism*, *immersion* and *social richness*) relate to Psi, which is one of the two contributing factors of presence outlined by Slater (2009). The second contributing factor of presence according to Slater (2009) is Pl, which relates to the '*it is here*' from of *transportation*, which is another conceptualisation of presence as outlined Lombard & Ditton (1997). However, this thesis goes further by relating the two contributing factors of presence (i.e. Psi and Pl) outlined by Slater (2009) to the two contributing factors of corporeal presence outlined by Holz *et al.* (2011) (i.e. *representation* and *geometric correspondence*). As such these two contributing factors of corporeal presence can be linked through Psi and Pl to the same four previously mentioned conceptualisations of presence (i.e. *realism*, *immersion*, *social richness* and *transportation*), this link was first discussed in Chapter 1 and illustrated in Figure 1.3.

Bearing in mind what was just discussed the remainder of this chapter will have the following structure. The next section (see Section 2.1) introduces presence and corporeal presence, specifying the difference in relation to avatar UI. There will be further explanation on the linkage of corporeal presence through Psi and Pl to the four conceptualisation of presence outlined above. In Section 2.1.1 a discussion into the conceptualisation of presence as realism takes place, followed by Section 2.1.2 on immersion, Section 2.1.3 on social richness and Section 2.1.4 on transportation. Then the causes and effects of presence are outlined in relation to the four conceptualisations detailed above in Section 2.1.5. Then as the empirical studies (see Chapters 4, 5 and 6) are weighted heavily towards quantitative data gathered from participants during *gaze perception* trials, Section 2.2 introduces gaze perception and the *Mona Lisa Effect* is discussed in Section 2.3 as it relates to the experimental set-up also. The chapter comes to a close with a brief summary of the background literature and an introduction to the next chapter in Section 2.4.

2.1 Presence

This section introduces the concepts of *presence*, relating presence to avatar UI research and the scope of this thesis, increasing a user's perceived presence for the rendered avatar on a 2D display. Heerink *et al.* (2010) discuss presence (or *social* presence) in two distinct scenarios: (1) in terms of virtual reality (VR) where the user feels present in the virtual environment, commonly defined as the sense of *being there* (Witmer & Singer, 1998); (2) the feeling of being in the company of another social entity, such as an avatar UI. Research by (Norman, 2007) indicates that humans are likely to instinctively treat technical devices (i.e. computers, phones and cars) as social beings, it does not seem like too much of a leap to assume they will do the same for an avatar UI. Especially as the avatar UI will convey a sense of presence, according to Reeves & Nass (1996). The second scenario of presence put forth by Heerink *et al.* (2010) is important to this thesis, because in order to be a successful social agent, an avatar will not only need to be present in the user's real-world environment but in fact make their presence felt by the user they wish to interact with.

Biocca *et al.* (2003) define the more social aspect of presence under three categories: *co-presence*, *co-location* and *mutual awareness*. All three can relate to both scenarios put forth by Heerink *et al.* (2010) to a greater or lesser extent, as all three categories do express a form of *togetherness*. The *co-presence* category relates to the sensory awareness of an embodied other. The *co-location* category is based around the experience of psychological involvement, including the concepts of *saliency*, *immediacy*, *intimacy* and to make one's self known. The *mutual awareness* category refers to the behavioural interaction which relies more so on immediacy behaviours through which social richness establishes presence (Lombard & Ditton, 1997; Rettie, 2003). Lombard *et al.* (2000, p.1) define presence as '*the perceptual illusion of non-mediation*'. Lombard *et al.* (2000) use the term *perceptual* to indicate that this phenomenon involves continuous (i.e. real-time) responses to objects and entities in a person's environment through the human sensory, cognitive and affective processing systems, important aspects of the '*Theory of Mind*' (Perner, 1999). For an *illusion of non-mediation* to occur, Lombard *et al.* (2000) say a person must fail to perceive or acknowledge the existence of a medium (i.e. 2D display) in their immediate surroundings and consequently, the person will respond as if the medium was not there at all. Although there are many concepts of presence, Lombard & Ditton (1997) have outlined six of the core conceptualisations of presence which relate to this research, they are as follows:

- **Realism**: *Perceptual* or *social*
- **Immersion**: The sensation of being immersed in a mediated environment.
- **Social richness**: The possible level of *warmth* or *intimacy* experienced via a medium.
- **Transportation**: The experienced sensations that '*you are there*', '*it is here*' or '*we are together*' (i.e. the avatar and user occupy a *shared space*).

- **Social actor within medium**: A *para-social* interaction, commonly seen between a TV presenter and the audience of TV viewers.

- **Medium as social actor**: The treatment of computers or other inanimate objects as social entities.

However, the scope of this thesis is only concerned with first four conceptualisations (i.e. realism, immersion, social richness and transportation) of presence and how these four conceptualisations relate to Psi and Pl, the two contributing factors of presence (Slater, 2009) (see Figure 1.3). Once again, the scope of this thesis required a more focused analysis of presence in relation to avatar UI research. *Corporeal* presence as outlined by Holz *et al.* (2011), is a more specific form of presence relating to the perception of the avatar's body (i.e. representation) within the user's real-world environment and is less concerned with the avatar's social presence.

The two contributing factors of presence (i.e. Psi and Pl (Slater, 2009)) are similar to the two contributing factors of corporeal presence (see Section 2): (1) *Representation* - which is similar to Psi; (2) *Geometric Correspondence* - which is similar to Pl.

It is important to establish what is meant by the term presence in this thesis and especially the idea of how corporeal presence relates to presence through having similar contributing factors. Clearly defining the concept of presence as a whole as it relates directly to avatars interfaces and will form the basis for all potential solutions to the problem statement set out in the previous chapter (see Section 1.1). As avatars are social actors that appear through a medium (i.e. 2D display) to exert their existence in the user's environment, if they don't achieve a strong sense of presence on this medium the user may not feel the need to interact with them. In addition to this, any interaction that does occur may come off as being trivial due to a lack of connection between the user and the avatar. Once the six conceptualisations are each introduced and related to avatar UI research in this chapter, the *causes and effects* of presence (see Section 2.1.5) will be discussed from a visual sensory perspective as this relates more to the development of an avatar UI.

The next section discusses the concept of presence called *realism* (Lombard & Ditton, 1997).

2.1.1 Realism

In avatar UI research it is important to implement a virtual agent's avatar so it acts like a human, if it looks like a human. Much avatar UI research focuses on producing realistic behaviours that drive social agents, e.g. *personality*, *memory*, *gestures*, *facial expressions*, *speech/dialogue* and *realistic gaze* (Gebhard *et al.*, 2008; Jan *et al.*, 2009; Mumme *et al.*, 2009). Bosse *et al.* (2007) argue that such realism in the avatar's behaviours is important to successfully establish social agents as credible communicators. You can lower a user's expectations of human-like behaviour by using non-human entities as the representation of the avatar, such as the animated robot used in *MACK* (Cassell *et al.*, 2002). Holz *et al.* (2008) achieve the lowering of their users' expectations by using a *cartoon-like* avatar on their mixed reality (MR) robotic platform. In the field of robotics when a robot looks human-like but does not act human-like, they fall into what Mori (1970); Mori *et al.* (2012) coined as the *uncanny valley*. Just as this happens for a human-like robot in the physical world, it can also happen for a graphically rendered avatar in a virtual world. Holz *et al.* (2008) argue that realism is an obstacle that must be carefully navigated when developing a virtual agent's avatar or a robot with human-like attributes. Holz *et al.* (2008) state that a strong *anthropomorphic* archetype, while necessary in order to build upon the evocative power of these agents, does in fact only serve to increase a user's expectations in performance of the agent's human-like nature. Consequently, this increase in a user's expectations severely raises any behavioural complexity required for the practical applications of these avatars or robots.

As the level of realism in an avatar's appearance approaches a human-like level, which is sometimes referred to photo-realistic, the avatar's actions and behaviours must match the realism of their human-like appearance. If the avatar's actions and behaviours (e.g. gaze) do not match their human-like appearance, the user may experience uncanny valley when interacting with the avatar, which could potentially diminish the '*illusion of life*' (Thomas & Johnston, 1995) otherwise created by the avatar. A good example of this happening in 3D animation can be seen in the film '*The Adventures of Tintin: Secret of the Unicorn*' (Steven Spielberg, 2011) where on its release the term uncanny valley became widely known and is discussed in an on-line review[1] of the film how "*Tintin looks simultaneously too human and not human at all, his face weirdly [sic] fetal, his eyes glassy and vacant instead of bursting with animated life*".

Interestingly, Riek *et al.* (2009, p.5) state "*people are more empathetic toward human-like robots and less empathetic toward mechanical-looking robots*" in their research. Riek *et al.* (2009) discuss their results as being compatible with '*Simulation

[1] The Biggest Problem With the Tintin Movie Might Be Tintin Himself: http://www.vulture.com/2011/07/the_biggest_problem_with_the_t.html

Theory' ((Goldman, 2008) in (Riek *et al.*, 2009)), when people mentally *simulate* the situation of other agents in order to understand their mental and emotive state. The more similar the other agent is to the '*empathizer*' the stronger the empathy process is. Also, Tao (2009) discusses how realism leads to the believability of an agent and is crucial to the creation of empathy. Similarly, Baylor & Kim (2005); Baylor (2005) found that users perceived more realistic looking agents as more believable instructors in their experiments. An avatar would be more successful if they had a strong resemblance to the human user, in terms of behaviour and appearance, in order for the user to establish rapport with the avatar. Mimicking another's body language is often used to establish rapport and having an avatar that strongly resembles a human would obviously allow for mimicking body language to occur. Other work (Gratch *et al.*, 2006, 2007; Wang & Gratch, 2009) emphasises the importance of mimicry (i.e. facial expressions, body language, speech inflections, etc.) to establish rapport and eventually trust, which is the foundation to any long term relationship. Furthermore, Bates *et al.* (1994) argue that believability will never arise from copying reality directly and that mimicry is a necessity. Bates *et al.* (1994) elaborate on this point by making the analogy that an artist uses reality in the service of realism by carefully studying nature but never elevating realism above their fundamental goal.

Maintaining the 3D illusion of the avatar so that it correctly corresponds geometrically to its real-world surroundings from the user's perspective is important. It is the 3D illusion effect that allows the user to accurately interpret the avatar's gaze direction when the avatar is referencing PoI with gaze alone. Kipp & Gebhard (2008) are concerned with both the tracking and the rendering areas discussed previously. The *iGaze* system developed by Kipp & Gebhard (2008) takes the same approach and updates the rendering of the avatar to match the user's perspective at all times. This approach allows for more accurate and realistic gaze behaviours. Kipp & Gebhard (2008) show in their results that participants found that the continued updating of the 3D scene with the avatar from their own perspective was stimulating rather than uncomfortable and that the avatar's gaze behaviours, such as *dominance* or *submissiveness*, were perceived as such. Kipp & Gebhard (2008) conclude that the *iGaze* system with the constant head-tracking not only achieves a more realistic looking view of the 3D scene by sustaining the 3D illusion but also creating *immersion* (see Section 2.1.2) for the participants. The *Virtual Anatomy Assistant* (VAA) (Wiendl *et al.*, 2007) goes that much further by in incorporating visual occlusion, when the avatar appears behind a real-world physical object in the user's FOV, the avatar's representation is occluded by the object and from the user's perspective the avatar appears behind the object. This is also extended so the avatar casts realistic shadows onto real-world objects, greatly increasing the perceived *corporeal* presence of the VAA when augmented in the user's real-world surroundings.

Lombard & Ditton (1997) formulated realism as a concept of presence because of its intrinsic value to social agents and as an avatar UI is a visual interface, realism plays an important role. Lombard & Ditton (1997, p.5) state that realism is the "*degree to which a medium can produce seemingly accurate representations of objects, events, and people*". Thomas & Johnston (1995) discuss how the first process taught to newly employed animators at *Disney*, in the 1920s during the hand-drawn animation era, was how to animate a *sand bag* to display human characteristics, a process known as *anthropomorphism.* If the anthropomorphic version of the sand bag had enough human characteristics, viewers would be able to suspend their disbelief and empathise with the *sand bag*, creating the '*illusion of life*' (Thomas & Johnston, 1995).

The next section discusses the concept of presence called *immersion* (Lombard & Ditton, 1997) which falls into two distinct categories: *perceptual* and *psychological.*

2.1.2 Immersion

Lombard & Ditton (1997) conceptualise presence as immersion as two distinct categories:

1. *Perceptual* immersion refers to the degree in which the user feels they are submerged into a virtual environment (i.e. VR), where many of the user's senses are blocked to the outside world but stimulated with other sensory data to cause *psychophysical* responses. The term *psychophysical* comes from a branch of psychology, known as *psychophysics*, which is concerned with the quantitative relations between physical stimuli and their psychological effects. McQuiggan *et al.* (2008) state that *immersion* generally refers to the extent and nature of technology-provided sensory stimuli.

2. *Psychological* immersion results from the user feeling mentally involved, absorbed, engaged or engrossed in a place, object or person.

Seah & Cairns (2008) illustrate the differences of perceptual versus psychological with the example of the computer game *Tetris*[1], which does not offer a player an

[1] *Tetris* game: `http://en.wikipedia.org/wiki/Tetris`

opportunity to feel perceptual immersion, as there is no virtual environment in which the players can be present. Yet Tetris can provide psychologically immersion through involved game-play. It is important to discuss *involvement* at this point as it refers to the degree of attention and meaning assigned to a stimulant or a combination of stimuli. In the case of an avatar UI the stimuli would be visual on a 2D display and audio, much like a computer game (e.g. Tetris). Peters *et al.* (2009) describes engaging situations when playing computer games, as the feeling of losing oneself in the world of the game, oblivious to the things happening outside of that world. Involvement leads to psychological immersion; immersion from heavy involvement can lead to a feeling of losing track of the passage of time. People who become completely engrossed in a book often refer to a similar feeling of losing track of the passage of time due to high levels of involvement in the narrative and this '*distortion of temporal experience*' is common in a mental state referred to as *flow* (Nakamura & Csikszentmihalyi, 2009).

Giannachi *et al.* (2009); Pan *et al.* (2008); Zanbaka *et al.* (2007) developed systems that continuously updated the avatar to match the user's perspective and they achieved high levels of immersion. The avatars used in experiments by Giannachi *et al.* (2009); Zanbaka *et al.* (2007) did not engage participants in direct conversation. The results from Zanbaka *et al.* (2007) show that immersion did not enhance social influence but could actually have a negative impact on the user when they are being watched by the avatar, while doing a complex task. Pan *et al.* (2008) and Giannachi *et al.* (2009) both make use of the *CAVE* (Cruz-Neira *et al.*, 1993) system in their experimental set-ups, CAVE is a fully immersive environment (i.e. VR theatre) as can be seen in Figure 2.1.

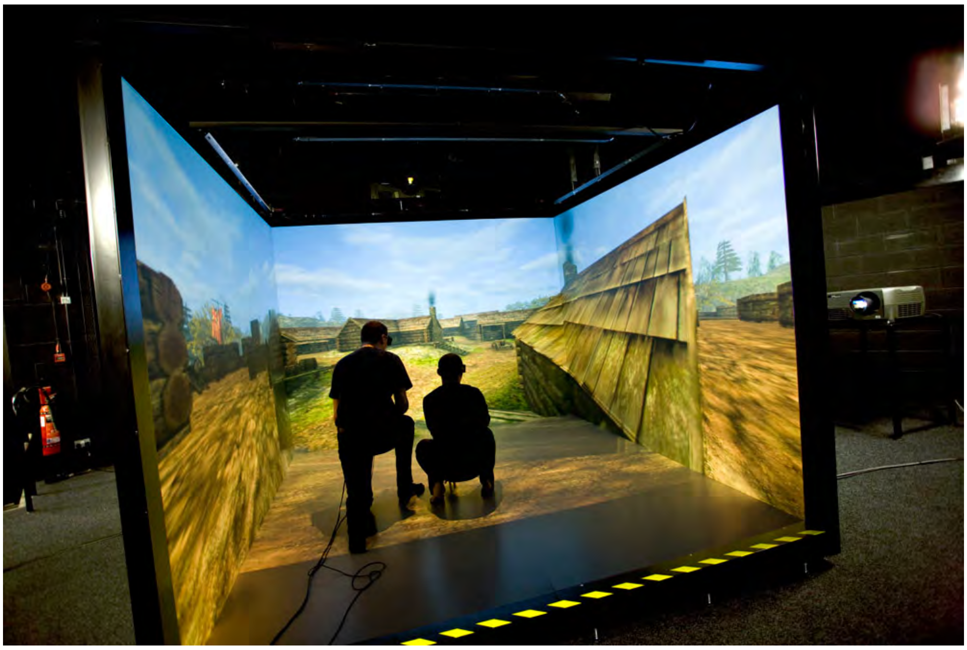


Figure 2.1: *CAVE* system in operation with two occupants.

The results of an experiment by Babu *et al.* (2007) show that the use of immersive avatars to help participants learn social protocols is significantly higher than when the same social protocols are learned from a written and illustrated guide. Babu *et al.* (2007) set out to teach human users social conversational protocols through the use of an avatar UI; their goal was to see if immersive virtual humans could act as instructors and teach users the verbal and non-verbal protocols of the Indian language better than any written and illustrated guide. Babu *et al.* (2007) and Zanbaka *et al.* (2007) both used real humans as instructors to act as the control condition for their respective experiments.

Kipp & Gebhard (2008) establish immersion with the *iGaze* system in two ways: (1) the use of a 3D illusionistic effect that makes the user feel as their movements are effecting the 3D image, just as if they were looking through a real-life window; (2) the constant updating of the agents gaze to follow the user's head position by establishing a well-known phenomenon that occurs in 2D images of people looking directly out of the image called the *Mona Lisa Effect*.

In contrast to semi-immersive systems such as *CAVE*, there are fully-immersive systems that have the user wear a head mounted display (HMD) in order to view an avatar UI (e.g. the *virtual autonomy assistant* (VAA) (Wiendl *et al.*, 2007)) by augmenting the avatar over the user's real-world surroundings. The use of a HMD does increase the avatar's perceived presence as the avatar is viewed stereoscopically, however, the wearing of such equipment is not practical in many scenarios.

The next section discusses a concept of presence called *social richness* (Lombard & Ditton, 1997), where the emphasis is on the establishment of *intimacy* and *immediacy* as they play important roles in successful communication.

2.1.3 Social Richness

Lombard & Ditton (1997) state that presence as *social richness* is related to two important concepts: *intimacy* and *immediacy*. These concepts were originally founded in the area of *non-mediated interpersonal communication* (Lombard & Ditton, 1997) discussed by Choi *et al.* (2001), which can be applied directly to social agents and their avatar representations.

Intimacy

Choi *et al.* (2001) state that the experience of intimacy is closely related to the expression of non-verbal involvement. Intimacy is the perception an entity has of another entity resulting in familiarity, through physical proximity, eye contact, intimacy of conversation topic, amount of smiling and other behaviours (Choi *et al.*, 2001). In a human context, intimate behaviours would surmount to a close or warm friendship. Lombard & Ditton (1997) clarify that an overall level of intimacy is reached between two entities when an equilibrium is achieved between the *approach* and *avoidance* forces, which were first discussed by Argyle & Dean (1965) (referenced by Lombard & Ditton (1997)), these are as follows:

- An *approach* force is a psychological reaction that happens when a person is compelled to move towards another person in order to gain intimacy during an interaction; both parties tend to move towards each other until a satisfactory state is reached.
- An *avoidance* force is the complete opposite of an *approach* force and is used when a person wants to prevent intimacy from occurring or reduce the current level of intimacy.

In a normal encounter people will always achieve an equilibrium between these forces, in order to achieve a satisfactory or acceptable level of intimacy, social conventions will often dictate the required levels. The list of intimacy behaviours include factors such as: posture and arm position, trunk and body orientation, gestures, facial expressions, body relaxation, touching, laughter, speech duration, voice quality, laughter, olfactory (i.e. sense of smell) cues and many others. According to Lombard & Ditton (1997) a social entity with a high level social richness will adjust more precisely to the overall level of intimacy required during any interactions with other social entities.

Intimacy can influence how users interact with human-like avatars. Bailenson *et al.* (2001) outline how users were just as unwilling to approach an avatar closer than necessary, just like approaching a real human for the first time, when the avatar displayed realistic intimacy behaviour. The user instinctively established a *social contract* with the social agents, this in turn helps maintain social richness during interactions. Bailenson *et al.* (2001) suggest in their first study how this behaviour may change if *self-identity* is used to manipulate the perceived level of intimacy. Bailenson *et al.* (2001) touch on the topic of *personal space* and how it correlates to intimacy. Their observations show how the size of a *personal space bubble* between two people is inversely proportional to the level of intimacy they feel towards each other.

Bailenson *et al.* (2001) were able to conclude that people who identify aspects of themselves in a social agent, experience with the social agent increases in the level of intimacy and as a result they are willing to reduce their *personal space bubble.* Pan *et al.* (2008) also looked into this social dynamic and their investigation centred on the interaction style between a virtual woman (i.e. a female avatar) and male participants.

Immediacy

Mehrabian (1981) states that *immediacy* is the sense of psychological closeness. However, the definitions of *immediacy* seems to change depending on the author, the concept of *immediacy* seem to revolve around the *here and now* and the importance of a current interaction. Lombard & Ditton (1997) argue that the use of language can create a sense of *psychological closeness* or immediacy, and therefore any medium that transmits language can vary the language in order to vary the immediacy. Stucky *et al.* (2009) discuss the idea of *co-presence* as the process of the user controlling an avatar (i.e. self-representation of the human user) in a virtual world and having non-mediated interactions within that environment, to such an extent that the medium of communication (i.e. the 2D display the user is viewing the game on) fades away.

Stucky *et al.* (2009) argue that *co-presence* is essential to socialising and a variety of learning approaches that rely on real-time, contextual interactions between avatars or what is essentially the human users the avatars represent. This idea does not have to be restricted to just human controlled avatars, it can be applied to avatars that communicate in a non-mediated way with human users in their real-world surroundings. Stucky *et al.* (2009) state the advantages of being able to direct focused based attention on a variety of communication cues: (1) *accountability from others to engage in the interaction*; (2) *immediacy of a response*; (3) *shared context for the conversation.*

The next section discusses the concept of presence referred to as *transportation* (Lombard & Ditton, 1997). Lombard & Ditton (1997) split this concept into three different scenarios: '*you are there*', '*it is here*' and '*we are together*'. The scenario that is central to the argument in this thesis is the second type of transportation, '*it is here*', as this is the process of making the user believe the avatar is in the user's real-world surroundings.

2.1.4 Transportation

In this section the concept of presence as *transportation* (Lombard & Ditton, 1997) is outlined. Lombard & Ditton (1997) split the concept of transportation in three distinct types:

- '*You are there*': The user is transported to another place (e.g. Computer Game Worlds or VR).

- '*It is here*': Another place and the objects within it are transported to the user (e.g. augmented reality (AR) or semi-immersive environments).

- '*We are together*': Two (or more) communicators are transported together to a place that they share (e.g. Multi-Player Games or Virtual Worlds like *Second Life* (SL)[1]).

All three types of transportation outlined above are effected to a greater or lesser degree by the three categories of *social* presence put forth by Biocca *et al.* (2003): *co-presence*, *co-location* and *mutual awareness* (see Section 2.1 for more detail). The concept of transportation is of benefit when it comes to an avatar UI as it what drives the felling of *togetherness* experienced by a user, as discussed by Biocca *et al.* (2003). The scope of interest for this thesis is only concerned with the '*it is here*' type of transportation and the next section will discuss it in more detail.

'*It is here*'

Lombard & Ditton (1997) define this type of transportation as the process of bringing objects, people or avatars to the user's location by displaying them through some sort of semi-immersive medium, such as a 2D displays (Babu *et al.*, 2007; Bailenson *et al.*, 2001; Cassell *et al.*, 2002; Kipp & Gebhard, 2008; McCauley & D'Mello, 2006), projectors in VR theatres (Giannachi *et al.*, 2009; Kipp & Gebhard, 2008; Kruppa *et al.*, 2005; Pan *et al.*, 2008) and HMDs with AR capabilities (Mac

[1] *Second Life (SL)* Virtual World: `http://bit.ly/SecondLifeVR`

Namee & Kelleher, 2009; Wiendl *et al.*, 2007; Zanbaka *et al.*, 2007). Lombard & Ditton (1997) state that it has been observed that some television viewers (i.e. like the '*you are there*' type of transportation) do not feel as if they are being taken into the television programme's world if the programme does not directly address them. If the '*it is here*' type of transportation is effective, an avatar should be able to correspond geometrically correct to the user's real-world surroundings from that user's perspective and as geometric correspondence (Holz *et al.*, 2011) is a contributing factor of corporeal presence. The '*it is here*' type of transportation plays a key role in the creation of corporeal presence for an avatar UI. Kruppa *et al.* (2005) track the user's movements via the user's position and orientation throughout an intelligent environment. Placing a projector on tracks attached to the ceiling, allows the projector to move around the room, which in turn enables the *Virtual Room Inhabitant* (VRI) to be projected to match the user's perspective onto any surface beside an object the user is interacting with. The ability of the projector to move on rails throughout the user's environment literally means the avatar is transported with you and completely fulfils the '*it is here*' type of transportation. VRI maintains the 3D illusion for the avatar as the user is constantly in the sweet spot for the 2D display and VRI as an avatar UI achieves the '*the perceptual illusion of non-mediation*' (Lombard *et al.*, 2000, p.1).

O'Hare *et al.* (2004) and Mac Namee & Kelleher (2009) (see Figure 2.2) use head mounted AR displays with front-facing cameras placed over both of the user's eyes to capture a live video stream from each of the user's eyes perspectives. These captured video streams are then rendered to the corresponding liquid crystal display (LCD) panels placed in front of each of the user's eyes. This then recreates a stereoscopic image of the user's environment from the user's perspective with the addition of 3D representations of an avatar augmented over each the video streams in accordance with the correct perspective for each eye. However, the 3D representations are usually anchored to a physical marker in the user's real-world surroundings, similar to the AR bar-codes visible in Figure 2.2. AR is a powerful method to create the sense of an avatar being transported to the user's real-world surroundings but it comes at a cost, both in terms of the need for expensive hardware and this hardware can be cumbersome or impractical to use or wear for most avatar UI scenarios. Such as the causal encounters a user may experience with an avatar in a museum or the lobby of a public building. Similar to *MACK* (Cassell *et al.*, 2002) and *MIKI* (McCauley & D'Mello, 2006), where a user would not be required to stand in a VR theatre or wear cumbersome head gear in order to interact with the avatar. Babu *et al.* (2007); Bailenson *et al.* (2001); Cassell *et al.* (2002); Kipp & Gebhard (2008); McCauley & D'Mello (2006) all take an approach where their respective systems conform to the '*it is here*' type of transportation where the avatar UI will engage the user from a semi-immersive display (e.g. 2D display, projectors, etc.) which could be placed throughout the user's environment.

The next section discusses the causes and effects of presence in the context of this thesis with clear focus on the visual characteristics of 2D displays that contribute to the creation of presence for an avatar.

2.1.5 Causes and Effects of Presence

Many senses (i.e. sight, sound, touch and smell) as well as one's own cognitive processes contribute to the creation of presence in terms of the four types (i.e.

Figure 2.2: **Left image**: The *Stepping off the Stage* (SOTS) agent (Mac Namee & Kelleher, 2009) is on a moving robotic platform and is being followed by a user wearing an AR capable HMD. **Right image**: The SOTS agent having stepped off the *stage* (i.e. robotic platform) is on a desk and describing a printer. AR bar-codes are visible in both images and are needed by the AR system to augment the 3D animated rabbit (i.e. SOTS agent's avatar) over the real-world scene in the user's FOV, which essentially grounds the SOTS agent to locations.

realism, immersion, social richness and transportation) discussed in the previous four sections. However, the scope of our interest is only concerned with the visual aspect for the creation of presence for an avatar UI and specifically, how to graphically render an avatar onto a 2D display in order to increase a user's perception of presence for the avatar. As a result of this the original list of causes and effects by Lombard & Ditton (1997) have been shortened as follows:

- *Image quality*: The perceived quality of the image (i.e. rendering of the avatar and 3D scene it is surrounded by) can have a huge impact on the sense of presence the user will experience. Graphics hardware and 2D display technology play a large role in the production of high resolution images that can be perceived to be photo-realistic. An avatar UI that lacks in image quality can suffer from a loss in perceived presence as described by De Freitas *et al.* (2010) (see Section 2.1.2).

- *Image size*: Size does matter, in the context of avatar UI research and specifically when rendering life-like avatars to displays within a user's environment.

The preferable display size should be capable of rendering life-size representations of virtual agents, this will invoke a more natural and realistic response from the user (i.e. CAVE (Cruz-Neira *et al.*, 1993) system), especially if the virtual agent is human-like in appearance.

- *Camera techniques*: There is a variety of ways in which camera angles can be used to create a greater sense of presence. In the television programming domain, presenters usually address the viewers by looking down the camera lens and talking directly to them, a good example of this is a television news reader reading a news bulletin. This effect gives a real sense of immediacy to the interaction between the presenter and the viewer. The avatar will take this direct address approach when engaging with users, maintaining eye contact throughout any face-to-face conversation that take place. Examples of this can be seen in the research by Dohi *et al.* (2009); Gebhard *et al.* (2008); Kipp & Gebhard (2008).

- *Viewing distance*: Tracking a user's viewing angle is particularly useful for avatar UI research where the user will engage an avatar UI in a more one-to-one interaction style on a 2D display. Rendering the avatar in their correct *proportions* in relation to the user perspective in their FoV is also important in order to maintain a strong sense of visual presence. This characteristic relates to the concepts of realism (see Section 2.1.1) and transportation (see Section 2.1.4).

- *Dimensionality*: Lombard & Ditton (1997) describe this characteristic as the process of adding *depth* and *perspective* to 2D displays to give a 3D illusion.

A good archetype of this can be seen in the virtual window created by the Winscape (RationalCraft.com, 2010) project. The use of 3D display technology quickly turns a 2D display into a stereoscopic one, increasing the user's sense of presence by adding real perceptual dimensionality.

For the most part this thesis is concerned with the *image size*, *viewing distance*, *dimensionality* and *camera techniques* display characteristics for achieving presence. The next section details the literature relating to gaze interpretability with a core focus on *gaze perception* in Section 2.2.

2.2 Gaze Perception

As a key area of interest, *gaze perception* is the only means of communication evaluated during the three empirical studies. In physiological terms *gaze* refers to the coordinated movement of the eyes and neck in order to facilitate any number of gaze behaviours (Argyle *et al.*, 1973; Langton, 2000; Mirenda *et al.*, 1983). Gaze perception includes the four behaviours outlined by Poggi *et al.* (2000): (1) *seeing*: when the eyes are used strictly for vision; (2) *looking*: when eyes are directed with the intention of seeing; (3) thinking: letting others know you are thinking by closing of eyes or directing eyes up, assist the thinking process; (4) *talking*: communicating with the eyes (i.e. eye actions and movements) in order to communicate information. Not only is it important to use gaze behaviours during an interaction with another person but the ability to interpret the other person's gaze behaviour is also of great importance for face-to-face communication (Al Moubayed *et al.*, 2012; Beskow & Al Moubayed, 2010). The eyes are often referred to as the '*mirror to the mind*' allowing others to fully interpret a person's true meaning during an interaction: allowing us to be affected by others as well as affecting others ourselves and allowing us to direct attention alongside interpreting others' directed attention (Argyle & Cook, 1976; Argyle & Dean, 1965; Carpenter *et al.*, 2000; Kleinke, 1986). Monk & Gale (2002) break down gaze perception into three differing levels of *gaze awareness* as follows:

- *Mutual gaze awareness*: The awareness one person has when they know whether another person is looking directly at them, i.e. *eye contact*.
- *Partial gaze awareness*: The awareness one person has when they know the direction another person is looking (up, down, left, or right).
- *Full gaze awareness*: The awareness one person has when they know the current object another person is directing their gaze towards, i.e. visual attention.

However, a problem that can sometimes occur with avatar interfaces is their limitation at portraying genuine eye gaze behaviours. The human eye is adept at interpreting gaze behaviours and is quick at finding fault in unnatural or disingenuous eye gaze behaviours. Peters *et al.* (2005) discuss how after the initial intrigue and novelty a human user has while interacting with an avatar dissipates, the user may begin to notice inconsistencies and implausibility in the avatar's gaze behaviours. This in turn can lead to a sharp decline in the quality of the interaction or may be the cause of the interaction's premature termination by the user. Peters *et al.* (2005, p.5) surmise that the avatar's ability to engage with the user is dependent on that user's perception of attention. Peters *et al.* (2005) argue that attention primarily acts as the control process, orientating the onlooker's senses towards stimuli of relevance to the engagement. Attention allows the onlooker to show they are involved with a speaker or an object of discussion in order to allow enhanced perceptual processing to happen. In order to elicit attention from someone, it is common place to direct eye gaze straight at them (Poggi *et al.*, 2000). Peters *et al.* (2005) hold the view that showing or perceiving interest would help establish rapport and help develop a relationship. Consequently, their model has the avatar direct its attention towards the user in order to show an interest and begin to establish a relationship.

Todorovic (2006) states that gaze perception not only depends on the position of the irises of the onlooker's eyes but also on the orientation of the onlooker's head. Furthermore, if the onlooker is not looking directly out of the image and is looking in a different direction but still out towards the user, only *partial gaze* is perceived. That said in certain circumstances if the user is in the sweet spot (see Section 1.1), they may perceive a stronger sense of *partial gaze*. However, when the user moves outside of the sweet spot their ability to correctly interpret gaze direction is reduced and any established *partial gaze* can become strained in the context of the user's real-world surroundings. *Partial gaze* might not be a problem in the domain of a painting, but when a 3D animated avatar that is trying to interact with a user by referencing PoI in the user's real-world surroundings, it can become a problem quite quickly. The TSB technique (see Chapter 3) prevents *partial gaze* and in fact elevates what would otherwise be perceived as *partial gaze awareness* to *full gaze awareness*, by allowing the user to fully interpret with high accuracy the directional gaze of the 3D avatar rendered on the 2D display. Tan *et al.* (2010) discuss the importance of *gaze awareness* in collaborative tasks, such as conveying the focus of the remote user's attention (or lack thereof) and is an important part of establishing the *inter-subjectivity* required for effective communication.

In a study by Eichner *et al.* (2007) participants' behaviour was monitored, specifically their eye gaze direction, in order to gauge their interest while watching *attentive presentation [virtual] agents* discussing two different MP3 players in a showroom scenario. Eichner *et al.* (2007) carried out a between subject experiment design with two conditions, one with the virtual agents responding to the participants eye gaze direction and the second having the agents react according to pre-defined points during the interaction with a participant. What is interesting about Eichner *et al.* (2007) set-up is that agents' ability to determine if the participant is paying attention to the presentation and how the agents deal with re-engaging a distracted participant.

Many systems (Cuijpers *et al.*, 2010; Kipp & Gebhard, 2008) like the *attentive presentation agents* system (Eichner *et al.*, 2007) have experimental set-ups that limit the participant to sitting in the sweet spot in front of large 2D displays in order for the eye tracking and gaze detection to be carried out. Limiting the user's ability to move and requiring them to stand in the sweet spot for the 2D display can effect the user's ability to interpret the avatar gaze direction, if the user's perspective changes evenly slightly, by either moving to the left or right of the sweet spot, their ability to interpret the avatar's gaze direction can deteriorate. Unless it can be guaranteed that the user is in the sweet spot, there is not much point attempting to direct the avatar's eye gaze anywhere else in the user's real-world surroundings but down the virtual lens to ensure the illusion of eye contact, i.e. the *Mona Lisa Effect* (see Section 2.3).

The next section introduces the *Mona Lisa Effect*, which helps establish eye contact between the avatar and user. The *Mona Lisa Effect* is created by harnessing the previous outlined list of causes and effects (see Section 2.1.5) to a greater or lesser extent.

2.3 The Mona Lisa Effect

The *Mona Lisa Effect* is an illusion named after the famous painting by *Leonardo da Vinci* (1506) that depicts a woman called *Mona Lisa* looking straight out of the canvas and seemingly peering into the eyes of any onlooker (Todorovic, 2006), this painting can be seen in Figure 2.3. The painting is a well-known example and the inspiration for the term *Mona Lisa Effect*. The onlookers of the painting often describe the sensation of Mona Lisa's eyes following them as they moved in front of the painting. The illusion also occurs on 2D displays where the rendered image of a person is looking straight out at the viewer, down the virtual lens in the 3D scene. The illusion becomes apparent as the gaze of the person rendered on the 2D display seems to follow the viewer as they move in front of the 2D display, regardless of their viewing angle and has being extensively studied by Koenderink *et al.* (2004).

However, when the onlooker is viewing an image of a person from outside the sweet spot their interpretation of the gaze direction of the subject within the image could be altered. In a thesis by Pol *et al.* (2009) that investigated if Mona Lisa Effect illusion held true regardless of the slant of the image from the user's perspective and their null hypothesis was that the slant of the image has no effect on how the user perceives the eye contact from the person within the image. Pol *et al.* (2009)

Figure 2.3: *Mona Lisa* a painting by Leonardo da Vinci that clearly illustrates the Mona Lisa Effect as described by Todorovic (2006), as the illusion where the eye gaze of the subject in the painting seems to follow the viewer no matter where the viewer stands in front of the painting.

discovered that the Mona Lisa Effect was not that straight forward and that the change in slant of the image (or alternatively the onlooker's viewing angle) has an effect on the perception of eye contact. They state there is a threshold of 20° to -20° from the sweet spot where the perception of eye contact remains strong. Pol *et al.* (2009) state that as viewing angles become greater than 60° to -60° there is a dramatic loss in perception of eye contact as at these angles the image will begin to suffer from lateral foreshortening. This is contrary to (Todorovic, 2006) prior beliefs that eye contact perception is not effected by slant. Cuijpers *et al.* (2010) carried out an experiment on fifteen participants where each participant had to rotate the image of a person on a 2D display who was looking directly down the virtual camera lens (i.e. creating the *Mona Lisa Effect*) and engaging them in direct eye contact until they could no longer perceive eye contact from the person displayed in the image. Cuijpers *et al.* (2010) proved that viewing angles of 60° to -60°, which are discussed above, are the thresholds at which eye contact is no longer perceived from viewing angles outside of these thresholds and these findings further support those by Pol *et al.* (2009).

Building on the results of their first experiment, Cuijpers *et al.* (2010) carried out a second experiment that was designed to test a participant's ability to interpret the gaze perception of a person looking out of a 2D display into the participant's real-world surroundings and not just looking down the virtual lens in order to make eye contact by engaging the Mona Lisa Effect. In their second experiment fifteen participants had to use their gaze perception to interpret the gaze direction of the person looking out of the 2D display into the participants real-world surroundings by holding a marker in a position anywhere along their perceived line of sight for the person looking out of the 2D display. The slant of the image varied, i.e. the user was not always in the sweet spot of the 2D display, as the experiment progressed. The results of this experiment showed that the participants perceived the person's gaze direction consistently wrong by overestimating the actual gaze direction by a multiple of two and in turn this confirms that using a standard 2D display while attempting to direct an avatar gaze anywhere other than down the virtual lens leads to inaccuracy, especially the further outside of the sweet spot for the 2D display the user is. Cuijpers *et al.* (2010) states that this overestimation for interpreting gaze is larger because people judge the sclera (i.e. is the white outer layer of the eyeball) of the eye as been the same in appearance regardless of slant of the image and this leads to overestimations.

In 3D computer graphics and particularly avatar UI, when the avatar looks down the lens of the virtual camera in a 3D scene it initiates the Mona Lisa Effect.

In terms of avatar UI research, the Mona Lisa Effect helps generate realistic eye contact between an avatar and the user viewing the avatar. Kipp & Gebhard (2008) harness this effect in their research to increase the realism of their avatar's gaze behaviours while it interacts with a subject during an interview and allows the avatar to portrait dominant and submissive gaze behaviours more acutely. This is especially important, as eye contact is a gaze behaviour that indicates the willingness of one social agent to engage in social interactions with another (Carpenter *et al.*, 2000; Peters *et al.*, 2005). Situations when the avatar is looking anywhere else other than down the virtual lens to make eye contact, it is extremely difficult for a viewer to interpret where the avatar is actually directing its gaze (Cuijpers *et al.*, 2010). For instance, if the avatar was to look to the left of the viewer, i.e. to the left of the virtual camera lens in the 3D scene, the viewer will always feel like the avatar is looking to their right regardless of where the viewer stands in front of the display. If the viewer is standing in the sweet spot their chances of interpreting the avatar's gaze direction increases dramatically, to the point where the viewer could guess what the avatar is looking at but only while the viewer remains stationary in the sweet spot. The Mona Lisa Effect has been shown to inhibit the interpretation of eye gaze direction when participants are viewing virtual agents on 2D display, as the avatar directs its gaze away from the virtual camera to disengage in eye contact and direct attention elsewhere, the user's has a false belief in their ability to correctly interpret the avatar's gaze due to their prior experience of realistic eye contact with the avatar. On the other hand when the avatar engages eye contact and initialises the Mona Lisa Effect, the illusion holds true for one or many onlookers (Al Moubayed & Skantze, 2011; Al Moubayed *et al.*, 2011). The next section summaries the background literature chapter used to support this thesis.

2.4 Summary of Background Literature

This chapter has introduced the literature on the Mona Lisa Effect, gaze perception and the concept of presence and how it relates to this thesis and the problem domain outlined in Section 1.1. To surmise the last of the three key areas discussed in the background literature, the Mona Lisa Effect is an important illusion that can easily be harnessed in order to help establish realistic eye contact between a human user and an avatar as it appears on a 2D display. Kipp & Gebhard (2008) found it was useful in their experiment to harness the Mona Lisa Effect in order to create a dominant presence for the rendered avatar when the avatar peered directly at the participants. However, in the context of the study by Kipp & Gebhard (2008) participants felt uncomfortable when they were gazed at in a dominant manner by avatar as the avatar made direct eye contact with the participants during the experiment. This feeling of discomfort is regardless of the fact that establishing eye contact is an important step for initiating a social interaction with another social entity (Peters *et al.*, 2005; Poggi *et al.*, 2000) but doing it in a dominant manner can be counter productive.

However, as stated in the literature, the Mona Lisa Effect has some major limitations, one being that at extreme viewing angles (i.e. image slant) the illusion is diminished or even broken completely, and with standard 2D display the Mona Lisa Effect can only really help the avatar make direct eye contact with the user, if the user is in or close to the sweet spot. If the avatar looks to the left or right of this straight out direction, the user will have difficultly interpreting the avatar's gaze direction and this difficulty increases the further from the sweet spot the user is viewing the 2D display. Therefore unless it can be guaranteed that the user will remain in the sweet spot for the entire interaction, it is not worth attempting to direct the user's attention to the avatar's gaze direction when it is referencing an object in the user's real-world surroundings. As Cuijpers *et al.* (2010) showed in their results, it is difficult for the user to interpret the avatar's gaze direction when they are outside of the sweet spot and users consistently overestimated the avatar's gaze direction. The user's overestimation increases with the acuteness of their viewing angle to the 2D display or as the slant of the image increases. In essence, this overestimation means a decrease in interpretation of the avatar's gaze direction and the overestimation becomes more exaggerated the further the avatar looks away from the straight out direction it takes when making realistic eye contact. If the user is not in the sweet spot for the 2D display, their ability to interpret the avatar's gaze direction decreases.

In terms of gaze perception the background literature details its importance to human communication and thus, proves it is worthwhile to ensure it is correct in human-to-avatar communication. As previously stated, the ability to interpret another social entity's gaze is important, as it is how humans engage one another in face-to-face interactions. One interlocutor (i.e. a person who takes part in a dialogue or conversation) directs their gaze towards a second interlocutor to initially get the second interlocutor's attention and also to show their willingness to engage in a social interaction with the second interlocutor (Peters *et al.*, 2005; Poggi *et al.*, 2000).

Previous research (Cuijpers *et al.*, 2010; Eichner *et al.*, 2007; Kipp & Gebhard, 2008) has harnessed human gaze perception already but has limited their participants to remaining in the sweet spot of the 2D display during any interaction with the avatar. What this means is that in a more natural interaction style, when a user interacts with an avatar UI or when the avatar UI attempts to engage the user in order to engage in an interaction, if the user is outside of the sweet spot of the 2D display, the user's gaze perception of the avatar's gaze behaviours will most likely be interpreted incorrectly or not at all.

The interpreting of an avatar's gaze by a user would not be possible if the user did not perceive the avatar's corporeal presence in the user's real-world surroundings. Ensuring that the 3D illusion of the avatar is maintained on the 2D display from the user's perspective at all times is key to establishing corporeal presence. Not only does it ensure that the user perceives the avatar's corporeal presence, it means the avatar correctly corresponds geometrically to the user's real-world surroundings and this in turn re-enforces the perceived corporeal presence, but also allowing for accurate interpretation of the avatar's gaze direction by the user as the avatar looks out of the 2D display and into the user's real-world surroundings. This thesis proposes to develop a graphical approach for rendering the avatar to a 2D display while matching the user's perspective as they move freely in front of the 2D display and thus, not limiting the user to the sweet spot during their interactions. The next chapter (see Chapter 3) introduces the TSB technique.

Chapter 3

The TSB Technique

To briefly reiterate the problem domain (see Section 1.1), it is the diminishing of the 3D illusion of an avatar on a standard 2D display from the user's perspective that occurs as the user moves away from the sweet spot of the 2D display. The avatar's 3D illusion diminishes on the 2D display because of lateral foreshortening, which means that the visual representation of the avatar is distorted from the user's perspective. A visual representation of this in effect can be seen in Figure 3.1. This has a knock on effect on the user's ability to interpret the avatar's gaze direction towards PoI in the user's real-world surroundings. From the user's perspective the avatar no longer achieves geometric correspondence between its virtual sub-space (i.e. the avatar's virtual projection rendered to the 2D display) and the physical sub-space (i.e. the user's real-world surroundings). As a result the user's perceived corporeal presence of the avatar is reduced. Lateral foreshortening not only limits the interpretability of the avatar's gaze direction, it also effects the avatar's visual representation that plays a role in creating the illusion that the avatar co-habits the user's real-world surroundings. As the illusion of co-habitation created by the avatar's visual representation decreases from the user's perspective, the perception of the avatar existing in their vicinity also decreases. The novelty of the TSB Technique detailed in this chapter is how it combats the above problem through a combination of three graphical processes:

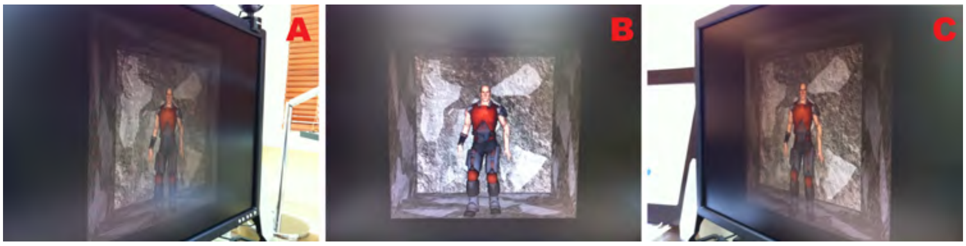


Figure 3.1: A visual representation showing how the 3D illusion deteriorates on a 2D display as the viewing angle veers away from the sweet spot: to the left in **A** and to the right in **C**. This series of three images is from the viewer's perspective. **A**: Acute viewing angle from the left-hand side of the 2D display shows deterioration of the 3D illusion caused by lateral foreshortening. **B**: Acute viewing angle from the right-hand side of the 2D display shows deterioration of the 3D illusion caused by lateral foreshortening. **C**: View of the 2D display from the sweet spot shows a 3D image that is correct for the tracked user's perspective, i.e. the optimal 90° viewing angle.

1. *Turning* (see Section 3.1.1) uses head-tracking data gathered from the user via the Microsoft Kinect to rotate the 3D scene that contains the avatar. This ensures that the user sees the 3D scene in the correct perspective from their current viewing angle and as this viewing angle changes, the 3D scene is updated accordingly. If the avatar wants to engage the user in eye contact all the avatar has to do is look down the virtual lens in the 3D scene and this initiates the Mona Lisa Effect (see Section 2.3).

2. *Stretching* (see Section 3.1.2) uses the same head-tracking data as the turning process to combat the image distortion caused by lateral foreshortening, which

appears from the user's perspective as they move outside of the sweet spot. It does this by stretching the rendered image on the 2D display and the amount of stretching applied is dependent on the viewing angle of the user. This stretching process ensures that the user sees a proportionally correct image of the 3D scene and avatar from their current viewing angle.

3. *Boxing* (see Section 3.1.3) encapsulates the avatar in a 3D box within the 3D scene. The 3D box is part of the 3D scene and as such is alerted by the two previous processes according to the user's current viewing angle. The function of the box is to limit the user's viewing angle to a realistic Field of View (FoV), as the 2D display has limited width and the image cannot be stretched indefinitely to compensate for the acute viewing angle of the user as they move further away from the sweet spot.

The combined effect allows for increased interpretability of the avatar gaze direction, consequently increasing the perceived corporeal presence for the avatar. In turn this increases the user's sense that the avatar actually exists in the user's real-world surroundings. The claims of the TSB technique (see Section 1.2 for more details) are listed as follows:

1. *Increase interpretability of the avatar's gaze direction*

2. *Increase perception of corporeal presence for the avatar*

The next three points further link the background literature previously seen in Chapter 2 to this research, stating exactly how *presence* will be achieved by each the three processes of the TSB technique:

Figure 3.2: The full TSB Technique in action. **A.** Viewed from far left. **B.** Viewed from sweet spot. **C.** Viewed from far right. (If viewing this figure on a computer monitor the viewer just needs to move to the left of the monitor for image **A** and to the right for image **C** in order to get a sense of how the 3D illusion is maintained from acute viewing angles by the TSB technique.

Previously outlined in the discussion on the causes and effects of presence (see Section 2.1.5), *camera techniques* are a display characteristic and correspond directly to the *turning* process in the TSB technique. As the avatar turns to look directly at the virtual camera inside a virtual scene, it creates realistic eye contact. The virtual camera's position in the virtual scene will continuously be updated to reflect the head position of the user being tracked in the real-world. Hence, the user's perspective of the 3D scene will be different, but in order for the avatar to engage in eye contact it has to look towards the virtual camera. This process helps establish the Mona Lisa Effect and increase the engagement felt by the user during interactions with the avatar due to the illusion of realistic eye contact from the avatar. This process also ensures a strong level of engagement felt by the user towards the avatar during any interaction, and might also lead to a high level of psychological immersion.

'*Dimensionality*' (see Section 2.1.5) corresponds to the *boxing* process as the avatar will be placed in a 3D room that has depth from the user's perspective and is similar to the RationalCraft.com (2010) project. The rendering of the 3D room will be continuously updated in relation to the user's viewing angle in order to maintain the effect of the user looking through a window into the avatar's world or the avatar looking out into the user real-world. This process should achieve the quasi perceptual immersion discussed previously.

'*Image Size*' and '*Viewing Distance*' (see Section 2.1.5) correspond somewhat to the *stretching* process in the TSB technique, when the user's *viewing angle* is considered to be part of this characteristic. The 3D scene with the avatar placed within it will be rendered depending on the viewing angle of the user (i.e. the user's look direction), which may change over time as they move in front of the display that the avatar is currently on. The image being rendered to that display will be stretched to compensate for lateral foreshortening, which is caused due to the user being at a viewing angle greater or less than the optimal ninety degrees (90°-right angle). This will happen on both the vertical and horizontal planes in order to compensate for height differences between users, where the differences in height would be far more significant between an adult user and a child user. This process helps maintain the effects of the *turning* and *boxing* within the TSB technique and in doing so maintains the avatar's level of presence in the user's environment which should lead to better communication between the user and the avatar.

Following on from this analysis, Section 3.1 introduces the three graphical processes of the TSB technique in greater detail (see Sections 3.1.1, 3.1.2 and 3.1.3). Finally, this chapter concludes with Section 3.2 where the combined result of all three graphical processes of the TSB technique is presented.

3.1 TSB: The Framework

The TSB framework presented in this chapter is a combination of three graphical processes[1]: *turning*, *stretching* and *boxing*. When combined, these three graphical processes combat the problem of viewing a 2D display from outside of the sweet spot (see Section 1.1), as the user strays further from the sweet spot the 3D illusion for the avatar diminishes this can be seen in Figure 3.1. The combined effect from the three graphical processes is to deliver a constant 3D illusion of the avatar on the 2D display from the user's perspective. The maintained 3D illusion is similar to the user looking through a *real-life* window at the avatar. This means that when the user moves in front of the display, the 3D scene continuously updates to match the user's perspective. As a result of this the avatar is able to accurately reference PoI (e.g. objects, places and people) in the user's real-world surroundings through gaze direction alone.

The TSB technique is dependent on head-tracking data for the tracked user and this data is retrieved from the Kinect SDK[2] *skeletal tracking* data, more specifically from the joint labelled: '*JointID.Head*'. This head-tracking data is smoothed out using a built-in process within the Kinect SDK before being input into the TSB technique as X, Y and Z coordinates for the tracked user's head position. The next three sections (see Sections 3.1.1, 3.1.2 and 3.1.3) explain how the three (i.e. turning, stretching and boxing) graphical processes are applied using this head-tracking data

[1] There is a short video available on **YouTube.com** called *TSB Technique in Action*, which illustrates the three graphical processes of the TSB technique individually and their combined effect: `http://youtu.be/OWDMGoDH640`

[2] *Microsoft's Kinect for Windows (SDK)*: Kinect for Windows was released in February 2012 with a Beta released in July 2011. Website: `http://www.microsoft.com/en-us/kinectforwindows/`

gathered from the tracked user's head position relative to the 2D display on which the avatar UI is rendered. Figure 3.3 shows how the user is tracked while positioned in front of the 2D display inside or outside the sweet spot.

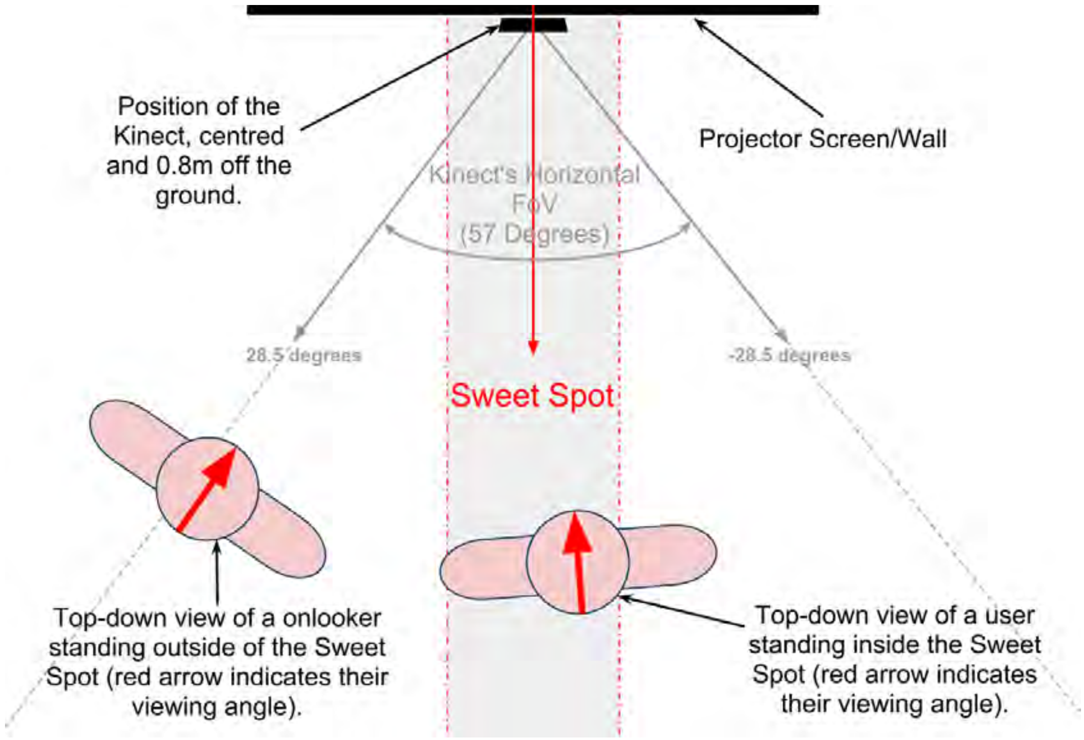


Figure 3.3: This diagram depicts a top-down view a 2D display, the Microsoft Kinect, a user in the sweet spot and a user outside of the sweet spot (i.e. the TSB technique only tracks one user at a time). There are two users depicted in the diagram, one illustrating a viewing position inside of the sweet spot (at an almost 90° viewing to the centre of the 2D display) and another illustrating a viewing position outside of the sweet spot. The large red arrows depicted on the two users' heads represent their viewing angles towards the 2D display. The further from the sweet spot a user views the rendered 3D image of the avatar on the 2d display, the more the 3D illusion of the avatar diminishes. This is caused by the increasingly obvious distortion of the 3D illusion due to lateral foreshortening as the user strays further from the *sweet spot*. The FoV for the TSB technique was dictated by the FoV of the Microsoft Kinect which is around 57° mark. The box of the boxing technique adds depth to the scene and grounds the avatar within a virtual space. From the user's perspective the box appears to be recessed back from the surface of the 2D display.

3.1.1 Turning

The turning effect is achieved by having a virtual camera's position in the 3D scene updated to match the user's head position in the real-world. When the user moves the 3D scene is updated to match their perspective, then when the avatar's gaze is directed towards the virtual camera in the 3D scene it subsequently seems to be directed towards the user and establishes the Mona Lisa Effect (see Section 2.3). When the avatar directs its gaze down the virtual camera lens it creates the illusion of eye contact between itself and the user regardless of the user's viewing angle. This illusion enables what seems to be realistic eye contact behaviour by the avatar toward the user from their perspective. It is true to say that any additional user(s) would also experience what they perceive to be consistent eye contact with the avatar due to the nature of the Mona Lisa Effect. However, the avatar is rendered to match the perspective of the tracked user's head position data, meaning that any additional user(s) would need to be sharing the current user's line of sight to the 2D display in order to fully appreciate the 3D illusion and interpret the eye contact as being realistic and meaningful (see Figure 3.4 **A**, **B** and **C** for an illustration of the *turning* process from three different perspectives). The addition of the next process, stretching, prevents the user from being restricted to the sweet spot area seen in Figure 1.4 and the user can move freely in front of the 2D display while still appreciating the rendered 3D illusion of the avatar.

Figure 3.4: **Turning only**: As the virtual camera moves, the avatar directs its gaze towards it. **A:** Viewed from far left. **B:** Viewed from sweet spot. **C:** Viewed from far right.

3.1.2 Stretching

The stretching process is achieved by normalising the same head-tracking data used in the turning process to calculate the user's viewing angle as a vector, this is then referred to as the '*eye position*'. Once the necessary adjustments are made to compensate for the position of the Kinect below the centre of the 2D display and after applying some smoothing to the raw data from the Kinect, two matrices are created using the eye position data: (1) view (or camera) matrix and (2) projection matrix. The view matrix uses the eye position data to create the virtual camera view from which the virtual camera will capture the 3D scene. The camera's view is continuously updated to match the actual tracked user's viewing angle through the eye position data. The projection matrix uses the inverse of the eye position data to create perspective in the 3D scene for the camera's view point. When both these matrices are applied to the 3D meshes (i.e. the avatar and the box) in the 3D scene they create a '*parallax effect*', where the objects in the distance appear to move faster than the objects closer to the camera in the rendered scene.

Normally a parallax effect occurs when viewing a 3D scene from within the 3D scene, however, in the case of the avatar UI the 3D scene is being viewed through a view-port (i.e. a virtual window) rendered to the 2D display. The further into the 3D scene a mesh is (i.e. the along the position Z axis) the more it is effected by the distortion of the changing perspective (i.e. the projection matrix) according to the user's current eye position and this causes the illusion of the view-port being stretched on the 2D display. This counteracts any distortion (e.g. narrowing or skewing) of the avatar's 3D image caused by lateral foreshortening when then user is at an acute viewing angle that would otherwise diminish the 3D illusion. Figure 3.5 illustrates the illusion of stretching in action, first when viewing an image from the sweet spot where the image looks stretched and second, from the viewing angle for which the image is rendered in the correct perspective. In Figure 3.6 **A**, **B** and **C** the avatar UI appears with the *stretching* process only switched on.

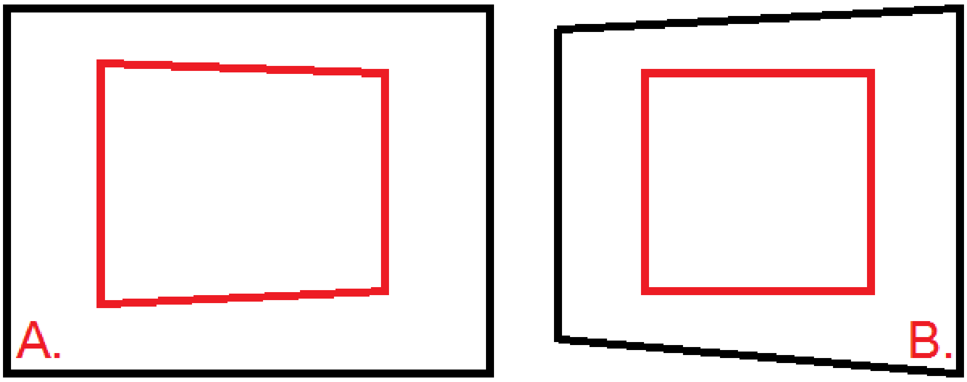


Figure 3.5: This is a simple illustration to help the reader understand the illusion of stretching that occurs. The black line represents the perimeter of the surface area of a 2D display and the red square represents the perimeter of a simple shape projected onto the 2D display. In part **A** the 2D display is being viewed from the sweet spot and the red square looks slightly stretched out of proportion on its left side, i.e. the vertical height on the left-hand side of the square is visibly taller than the right-hand side. In actuality the red square is being rendered to match a far right viewing angle. **B** illustrates what the user would see from a far right viewing angle and the red square appears to be square. The vertical height of the left-hand side of the 2D display appears to have gotten shorter in height. This shrinkage can be explained due to the fact that the user is now further away from the far left-hand side of the 2D display and the visual aspect of the 2D display tapers down when they are standing at this far right viewing position.

Figure 3.6: **Stretching only**: As the virtual camera moves the scene is stretched to compensate for lateral foreshortening. **A:** Viewed from far left. **B:** Viewed from sweet spot. **C:** Viewed from far right.

3.1.3 Boxing

This effect is achieved by placing the 3D avatar in a virtual box and from the user's perspective this box is recessed into the wall of the room in which they are standing. This is particularly effective if the avatar UI is projected directly onto the surface of a wall. Then as the user's perspective changes, the 3D scene is updated by the turning and stretching processes and the opening of the virtual box begins to behave like a window. In Figure 3.7 there is a series of three images, **A**, **B** and **C**, these images demonstrate the boxing process in effect. As previously seen in Figure 3.2 where a series of three images illustrates the TSB Technique in full effect, you can see how the *boxing* process looks from an extreme left (image **A**) and right (image **C**) viewing angles well outside of the sweet spot. The *boxing* effect contributes two important benefits to the overall TSB technique, they are:

1. The first benefit of the boxing process is creating realistic behaviour similar in effect to a view through a window in the real-world. The edges of the virtual box created by the boxing process occlude the avatar as the user moves to extreme viewing angles to the right or left sides of the 2D display. This creates a sensation of depth in the scene. Also, as the user moves closer to and further away from the surface of the 2D display, the scene is updated to reflect a realistic change in depth. The user can now see more (or less) of the internal walls of the virtual box depending on how close they are standing to the virtual box and their viewing angle. From the user's perspective the virtual box's opening immediately touches the 2D display surface from which it is then recessed to create the virtual window effect, acting more like a window frame.

This virtual window frame remains in place and does not move regardless of the user's viewing angle, remaining consistent in look and size as it sits at 0 on the Z axis. The stretching process has a more dramatic effect the further into the 3D virtual scene the rendering happens, along the Z axis.

2. The second benefit is maintaining a realistic boundary to the viewable area of the 2D display. As the 2D display has limitations as to how much space there is to display an image, there is no point warping an image to compensate for the perspective of the user, when the image becomes larger than the space available on the 2D display, i.e. the stretching process distorts the image of the avatar to a point where it can no longer fit on the 2D display. This would only become a problem at acute viewing angles. Adding the *boxing* effect prevents this from happening, thus preserving the integrity of the illusion.

Figure 3.7: **Boxing only**: As the user moves closer or further away from the display the scene is updated to reflect the change in depth from the user's perspective. **A:** Viewed from far away, the further the viewer is at the shallower the depth of the box. **B:** Viewer is closer the 2D display, more of the box's interior is becoming visible. **C:** Viewer is even closer to the 2D display and they have a clear view of the box's interior walls at this point due to increased depth.

3.2 The TSB Technique in Full Effect

The TSB technique uses a combination of the three graphical processes to render an avatar UI according to the tracked user's perspective onto a 2D display and increases the user's interpretability of the avatar's gaze direction towards PoI in the user's real-world surroundings. This increase in interpretability can in turn lead to a similar increase in the perceived corporeal presence of the avatar. Figure 3.2 shows three images, **A**, **B** and **C**, which demonstrate the TSB technique in action alleviating the problem of viewing 3D graphics on a 2D display from a non-optimal viewing angle. The TSB technique's ability to increase the perceived corporeal presence of the avatar should bolster the user's sense of psychological immersion, realism and social richness when interacting with the avatar maintaining a higher level of presence. This is especially true when the avatar is able to simulate realistic gaze behaviours.

The turning process specifically creates a high level of realistic gaze behaviours by harnessing the Mona Lisa Effect to deliver consistent eye contact with the user similar to how Kipp & Gebhard (2008) used this graphical process in their research. The turning process ensures the user's perspective of the avatar is correct by turning the 3D scene to match the user's viewing angle.

The stretching process literally stretches the rendered image of the avatar to help maintain the correct proportions of the image for the user's viewing angle counteracting the distortions caused by lateral foreshortening, further compliment-ing the turning process and sustaining the 3D illusion of the avatar from the user's perspective.

The boxing process adds depth to the 3D scene and the occlusion of the avatar by the frame of the virtual box as the user moves from the sweet spot to the far left or right adds realism to the 3D scene.

Furthermore, the benefits of the TSB technique's 3 graphical processes (i.e. turning, stretching and boxing) have a bearing on the user's perception of the avatar's corporeal presence and as such has a bearing on some of the conceptualisations of presence by Lombard & Ditton (1997). The turning process corresponds directly to *camera techniques* which formed part of the discussion on the causes and effects of presence (see Section 2.1.5). In the TSB technique the virtual camera's position within the virtual scene is continuously updated to reflect the head position of the user. Consequently, the avatar only needs to look directly at the virtual camera inside of the virtual scene to create the illusion of realistic eye contact between itself and the user. The turning process ensures the rendered scene matches the user's perspective as well as establishing the Mona Lisa Effect (see Section 2.3). In turn the Mona Lisa Effect could help to increase the psychological immersion felt by the user during any interaction with the avatar. However, testing this claim falls outside the scope of this thesis as it would require a specialised approach in order to analyse each participant's personal experience of psychological immersion, a very subjective topic and traditionally difficult to measure. This thesis is limited to analysis of the test results for each of the participants experience at interpreting the avatar's eye gaze direction, higher results indicating a higher level of perceived corporeal presence for the avatar. That being said, psychological immersion could be an interesting avenue to take for future work with the TSB technique.

The discussion of *dimensionality* in Section 2.1.5 corresponds to the boxing process of the TSB technique. This is due to the fact that the avatar will be placed in a virtual box that has depth from the user's perspective and the rendering of the virtual box will be continuously updated in relation to the user's head position in order to maintain the correct perspective for the tracked user's view through a virtual window at the avatar. In turn this dimensionality should help achieve realism and immersion for the user.

The next three chapters (i.e. Chapters 4, 5 and 6) outline each of the three empirical studies carried out to test the claims of the TSB technique (see Section 1.2). Chapter 4 details the empirical *Study 1* carried out in two parts, *Part I* evaluated the first claim and *Part II* evaluated the second claim of the TSB technique. *Part I* required each participant to do a series of a gaze perception trials where the avatar would look at PoI in the participant's real-world surroundings and the participant had to move to where they thought the avatar was directing its gaze. The quantitative data gathered from the participant was analysed to evaluate the first claim regarding increased interpretability of the avatar's gaze direction. This experiment was deigned to test for corporeal presence and relies on *quantitative* data Holz *et al.* (2011) due to its subjective nature, which makes it is hard to qualify. All the empirical studies in this thesis weigh heavily on quantifying participants behaviours through gaze perception trials. However, regardless of the difficulty with using surveys to qualify presence, *Part II* adapts questions from a standard presence survey questionnaire used by Witmer & Singer (1998) to supplement the quantitative data gathered. Chapter 5 details empirical *Study 2* which is a more elaborate gaze perception experiment that further tested the first claim of the TSB technique. Participants were require to guess where the avatar was directing its gaze in order to win a prize hidden under one of three upturned buckets. There was also a second part to the experiment where participants had to interpret the avatar's gaze direction as it looked past them towards PoI mounted on the surrounding walls.

Chapter 6 outlines empirical *Study 3* where participants carried out *Part I* of *Study 1* again, with the introduction of 3D display technology (i.e. anaglyph 3D), adding two new experimental conditions, *3D Control* and *3D TSB*. The data gathered was analysed in order to evaluate the first claim of the TSB technique.

Chapter 4

Empirical Study 1: '*42 Moves*'

This chapter presents the first in a series of three empirical studies carried out to evaluate the two claims of the TSB technique (see Section 1.2). This chapter is closely related to '*The Turning, Stretching and Boxing Technique: A Step in the Right Direction*' (Dunne *et al.*, 2012b) publication outlined in Section 1.4. This first empirical study evaluates both claims in relation to the benefits of the TSB technique when used in conjunction with an avatar UI on a 2D display, which are as follows:

1. **Increase interpretability of the avatar's gaze direction**: The TSB technique increases a user's accuracy at interpreting the avatar's gaze direction towards PoI in the user's real-world surroundings from the user's perspective. This increase is most noticeable when the user is interpreting the avatar's gaze direction from outside of the sweet spot for the 2D display.

2. **Increase perception of corporeal presence for the avatar**: The rendering of the 3D avatar is always correct from the tracked user's perspective, ensuring the visual fidelity of the rendered avatar as the user moves in front of the 2D display. In turn, this visual fidelity helps to establish the avatar's representation within the user's environment. Also, the visual fidelity of the rendered avatar means the avatar correctly corresponds geometrically to the user's environment from the user's perspective. Leading to the user's higher level of perceived corporeal presence for the avatar.

The '*42 Moves*' experiment tests for increases in perceived corporeal presence and is comprised of two parts. The first part of the experiment is based on measuring participants gaze perception and the second part is a survey questionnaire. The use of quantifiable measurements to gather data regarding the level of perceived corporeal presence is important (Holz *et al.*, 2011), as presence is subjective in nature and it is difficult to qualify for most people. Survey data alone would not suffice, therefore the first part of this experiment is weighted heavily towards the gathering of quantifiable data. Each participant had to interpret the gaze direction of the avatar as it looked towards one of seven floor markers (i.e. PoI in the user's real-world surroundings) forty-two times (i.e. once per trial). Figure 4.1 illustrates the arrangement of the seven floor markers in the '*42 Moves*' experiment laboratory. Each of these trials required the participant to stand on one of the seven floor markers and a participant can be seen standing on floor marker *7* in Figure 4.2 before he makes a guess as to which of the six remaining floor markers he perceives the avatar is directing its gaze towards. When they were happy with their guess they moved to stand on the guessed floor marker. This move was recorded and the outcome was either a correct or incorrect guess. The avatar never directed its gaze towards the floor marker on which the participant was standing at the beginning of the trial.

There were two experimental conditions, a control condition which displayed the avatar on a standard 2D display and a second experimental condition that used the TSB technique on the same standard 2D display. Each participant carried out the experiment once per experimental condition meaning they did all forty-two trials twice. When the TSB technique is used the combined effect of the three graphical processes detailed in Chapter 3 should enable participants to make more accurate interpretations of the avatar's gaze direction toward the floor markers, resulting in a higher accuracy rating. A comparison study of the gathered data for both of the experimental conditions from all the participants will highlight that the TSB technique outperformed the control. Also, geometric correspondence, which is a contributing factor of corporeal presence (Holz *et al.*, 2011) can be evaluated from the quantitative data gathered in part one of this experiment. Hence, a higher accuracy rating at interpreting the avatar's gaze direction would imply a higher level of geometric correspondence between the avatar and its real-world surroundings further implying a higher level of corporeal presence. Geometric correspondence can be directly related to Pl and as such both have a bearing on transportation, a conceptualisation of presence by Lombard & Ditton (1997).

The second part of this study required all participants to answer a survey questionnaire after both experimental conditions. The survey questions were adapted from a standard survey questionnaire outlined by Witmer & Singer (1998). The data gathered from the questionnaires helped to qualify the participants' perception

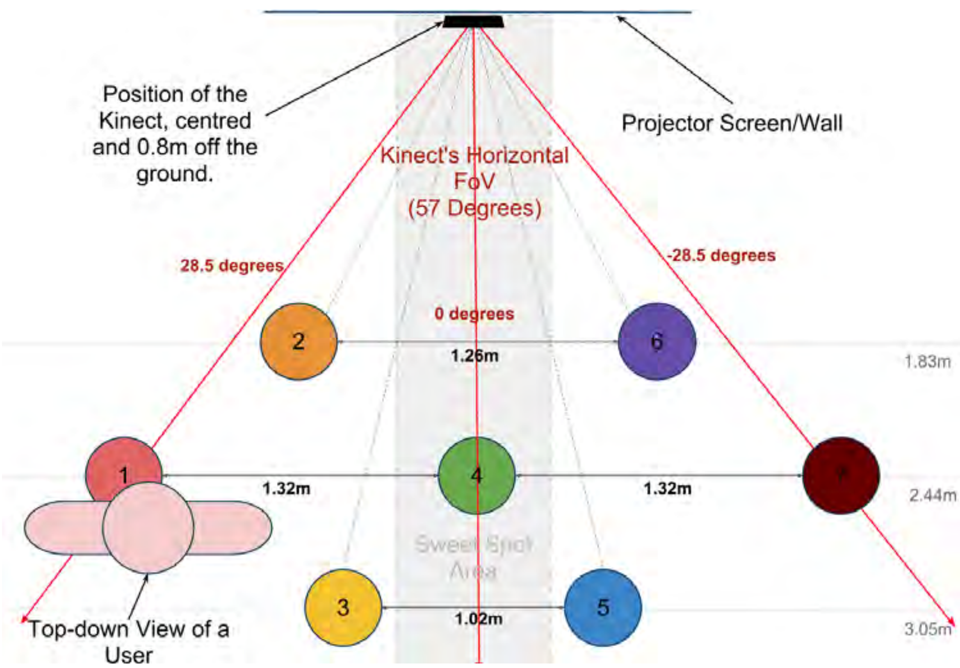


Figure 4.1: A floor plan for the '*42 Moves*' experiment. The floor markers and the sweet spot are visible.

Figure 4.2: A participant is standing on floor marker *7* waiting for their next move, for top-down view of experiment layout see Figure 4.1.

of the avatar's corporeal presence for both of the experimental conditions. Psi, the second contributing factor of corporeal presence according to Slater (2009) can be measured in the survey results, as it relates to the participants belief that '*what is happening is really happening*'. Psi, can also be related to the second contributing factor of corporeal presence: '*representation*', as outlined by Holz *et al.* (2011). Both Psi and representation have an impact on three of the six conceptualisations of presence as outlined by Lombard & Ditton (1997): realism (see Section 2.1.1), immersion (see Section 2.1.2) and social richness (see Section 2.1.3).

To briefly reiterate the problem first mentioned in Section 1.1, when the user is viewing the 2D display from outside the sweet spot the effectiveness of any 3D illusion is diminished due to lateral foreshortening distorting the user's perspective of the rendered image. The avatar's representation is distorted leading to loss in Psi and subsequently realism, immersion and social richness. As a result of the rendered avatar UI being distorted from the user's perspective, the avatar can no longer correspond geometrically correct to the user's real-world surroundings and this reduces the user's ability to use their own gaze perception to interpret the avatar's gaze direction.

In Section 4.1 the procedure used to carry out the first experiment is detailed. Section 4.1 also outlines the two experimental conditions used in more detail. Section 4.2 discusses the participants involved in the experiment and the reason for using such a diverse age range. Section 4.3 begins by reiterating that the focus of the survey was to evaluate the effect of the TSB technique on the corporeal presence perceived by the participants. This was achieved by examining the effect of the TSB technique on Psi, which relates directly to the avatar's representation, a contributing factor of corporeal presence (Holz *et al.*, 2011).

Next Section 4.4 details the results of the gaze interpretability part of the '*42 Moves*' experiment. This is followed by Section 4.5, which details the results from the survey portion of the experiment. Finally, this chapter is concluded with Section 4.7 where the conclusions for *Study 1* are examined.

4.1 Procedure

A *Wizard of Oz* experimental set-up (Dahlbäck *et al.*, 1993) was used so that the behaviours of the avatar could be fully simulated by the experimenter according to a pre-defined script. This was done in order to ensure that all participants had a consistent experience. The experiment took place in a large empty room with the avatar projected onto a wall. There were seven coloured circular markers placed on the floor that indicated specific positions in front of the projection. These can be seen in Figure 4.1 and again in Figure 4.2. During the experiment the avatar would direct its gaze towards one of the seven floor markers and the participant would have to guess which one the avatar was looking at. The participant's guesses were recorded with either a '*0*' for incorrect or a '*1*' for correct. Their final rating would be an indication of their accuracy at interpreting the avatar's gaze direction.

When each participant arrived to carry out their first experiment they were require to sign a release form (see Appendix A.1) and told they would be recorded on video while they carried out the experiment, then they were read a list of instructions by the experimenter. The instructions are listed below:

1. You will start on one of the seven coloured floor markers *(experimenter note: point to the coloured floor markers).*

2. A character will appear on the screen *(experimenter note: point to the screen)* in front of you, the character will not speak to you. He will only move his

head and eyes to look towards one of the remaining six floors marker you are not standing on *(experimenter note: stand in front of the screen facing the participant and move only your head and eyes to look at one of the floor markers).*

3. The character will never look at the floor marker you are currently standing on.

4. The character will look directly at you until you say you are ready, then the character will look at one of the other six floor markers.

5. You must remain stationary on your current floor marker and guess which of the other six floor markers the character is looking at. Once you move off your current floor marker the character will stop looking at one of the remaining six floor markers and look directly at you again.

6. Once you make a guess you are free to move to stand on the guessed floor marker and your guess will be recorded.

7. This process is repeated forty-two times, and generally takes less than twenty minutes but you can take as long as you need. If you need a break just ask.

8. Occasionally, a message will appear on the screen to ask you to move to another floor marker after you have made a guess. Once you have moved to that floor marker you may proceed as normal.

9. Finally, feel free to ask a question at any time during the experiment.

Each participant was then introduced to the character as it appeared on the 2D display in front of them, they were allowed to do a sample move as detailed above to familiarise themselves further with the process, this move was not recorded. Each participant was required to start the experiment by standing on one of the seven floor markers randomly assigned to them. Determined by the floor marker they started on, each participant proceeded through a pre-determined sequence of '*42 Moves*', twice. Once for both of the experimental conditions with at least a half-hour break between experimental conditions. The experimental conditions were as follows:

- **Control condition**: The avatar appears as it would on a regular 2D display, i.e. a projection onto a flat white surface. The rendering of the avatar does not update to reflect a participant's perspective. Hence, it increasingly suffers from lateral foreshortening as the participant's viewing angle becomes more acute than the sweet spot's optimal 90° viewing angle. Lateral foreshortening is at its worst when the participants were standing on floor markers, *1* and *7* (see Figure 4.1).

- **TSB condition**: The TSB technique is switched on, therefore the avatar's 3D rendering is continuously updated to reflect the participant's perspective, eliminating the distortion of the 3D illusion otherwise caused by lateral foreshortening. Meaning that the participant should perceive the full 3D illusion of the avatar no matter what floor marker they are standing on and this in turn should increase the participant's accuracy at interpreting the avatar's gaze direction.

To control for learning effects there were seven (i.e. one for each of the seven floor markers) pre-determined paths that were randomly generated to ensure each participants did all forty-two moves through the floor markers for each of the conditions, a different path for each condition. Meaning that regardless of how good a participant's memory was during their second trial they followed a completely different path and their previous answers gave them no advantage. Also, the order of the conditions (i.e. TSB condition or control condition) was varied between participants, half of the participants started with the *control condition* and the other half started with the *TSB condition*. Table 4.1 contains the entire sequence of moves for *Path 1* (see Appendix A.3 for complete sets of all seven paths). When a participant made a mistake the experimenter would display a message on the 2D display to indicate to the user that they were to move to a different floor marker. The message would remain on the display until the user had moved to the new floor marker. This ensured the participant could continue to carry out the pre-defined path. However, participants were not told that they had made a mistake, as this was would lead to negative re-enforcement and could potentially cause participants to feel dejected due to the fact that mistakes were common during the *control condition*. Each participant was guided through the path by the avatar's gaze only (i.e. eye, head and neck movements) and the avatar's gaze behaviour was exactly the same across both conditions. A video recording[1] of a participant carrying out the '*42 Moves*' experiment shows a side-by-side comparison of a participant progressing through both experimental conditions. Overlay graphics indicate the participant's accuracy at interpreting the avatar's gaze direction for both conditions as they progress through the experiment.

[1]Recording of the '*42 Moves*' Experiment showing a side-by-side comparison of the same participant carrying out both experimental conditions (the *Control* and *TSB* conditions): `http://youtu.be/R41C3xL0zfE`

Table 4.1: *Order* of the '*42 Moves*' in *Path 1*.

Order	Move	Order	Move	Order	Move	Order	Move
1	1-6	11	6-4	21	4-1	31	6-1
2	6-7	12	4-3	22	1-7	32	1-3
3	7-6	13	3-2	23	7-4	33	3-6
4	6-2	14	2-7	24	4-6	34	6-3
5	2-4	15	7-3	25	6-5	35	3-5
6	4-5	16	3-4	26	5-4	36	5-3
7	5-2	17	4-2	27	4-7	37	3-7
8	2-1	18	2-5	28	7-1	38	7-5
9	1-5	19	5-1	29	1-2	39	5-7
10	5-6	20	1-4	30	2-6	40	7-2
						41	2-3
						42	3-1

After each participant had conducted the forty-two moves required for a path to be complete, they had to answer a survey questionnaire; this was done for both experimental conditions.

4.2 Participants

There were thirty-one participants in total (nine females, twenty-two males) with ages ranging from six to sixty-four years. This diverse range in ages was selected as it is representative of the wide range of visitors to a museum or other large public building, which are the types of locations most likely to use large wall projected avatar UIs to engage with their visitors. Each participant signed a release form before carrying out the experiment (see Appendix A.1) that allowed for them to be video recorded during the experiment and the data gathered from them to be used in our research.

4.3 Survey

Participants took the same survey of six questions after each experimental condition; the results of the survey should indicate any increase in Psi experienced by the participant due to the effect of the TSB technique. Psi relates directly to the representation of the avatar in the participants' real-world surroundings and can be connected to the participants' experiences of three conceptualisations of presence (Lombard & Ditton, 1997) while interacting with the avatar: realism (see Section 2.1.1), immersion (see Section 2.1.2) and social richness (see Section 2.1.3). Holz *et al.* (2011) argues that surveys are not a good enough tool by themselves to measure a person's perception of corporeal presence for a MiRA (e.g. a representation of an avatar UI on a 2D display within a user's real-world surroundings) due to the subjective nature of corporeal presence. Surveys can only supplement quantitative data gathered during any experiment attempting to measure a participant's perception of corporeal presence. In the case of this empirical study both quantitative and qualitative data were gathered in order to measure a participant's perception of the avatar's corporeal presence.

The six question survey that supplements the '*42 Moves*' experiment in this empirical study was adapted from a standard presence survey questionnaire originally used by Witmer & Singer (1998). It was selected due to the ease at which Psi can be extrapolated from the questions in terms of the three conceptualisations of presence (i.e. realism, immersion and social richness) (Lombard & Ditton, 1997) outlined above. Each question was rated on a Likert scale as follows: 1: *Very low*, 2: *Low*, 3: *Average*, 4: *High* and 5: *Very high.* There were no open ended questions in this survey, it was felt that parsing any relevant data from opened ended answers about presence, which by its very nature is subjective, would of been unnecessary as the survey was only supplementing the quantitative data. The six questions were as follows:

1. *To what degree did you become so involved in doing the task that you lost all track of time?* The question rates the participant's level of immersion while interacting with the avatar UI. The higher the rating the higher a participant felt immersed during their experience interacting with the avatar.

2. *To what degree did you feel the 3D virtual character's head movements were natural?* The question addresses the participant's sense of realism for the avatar's head movements. A higher rating here would signify a higher sense of realistic movement or behaviour from the avatar when it was directing its gaze.

3. *To what degree did you feel the 3D virtual character's gaze direction towards the spots on the ground was realistic?* Like *Question 2*, this question is asking the participant to rate the realism of the avatar's directed gaze towards the floor markers, i.e. if the avatar was corresponding geometrically correct to the floor markers. A higher rating for this question can also indicate a higher level of Pl was experienced by the participant. Pl is primarily measured in the accuracy of interpreting the avatar's gaze direction in the first part of this study.

4. *To what degree did you feel the 3D virtual character was responsive to your actions?* This question is measuring the participant's sense of the avatar's social richness. Again, the higher the rating the higher the social richness, meaning the participant felt as if the avatar was reacting to their movements.

5. *To what degree did your experience with the 3D virtual character's gaze seem consistent with your real-world experiences?* This question is again attempting

to measure the participant's sense of the avatar's social richness by asking the participant to recall their previous experience at interpreting gaze direction and compare it to the avatar's ability to use gaze direction. It can also indicate the level of realism the participant senses from of the avatar's gaze direction.

6. *To what degree do you think the 3D virtual character was actually able to look out at the real-world 'spots' on the ground?* This question is a rewording of *Question 3* but is more concerned with social richness and realism of the avatar's behaviour than the realism it displays when reacting the participant's movements.

To confirm, the avatar's behaviour was exactly the same for both experimental conditions, therefore the null hypothesis would be that the participants' ratings remain similar across both experimental conditions. The next two sections split the results of the first empirical study into two parts. *Part I* (see Section 4.4) discusses the results of the gaze perception part of *Study 1* in regard to all the quantitative data gathered from participants. *Part II* (see Section 4.5) outlines the survey results, where the qualitative data gathered from the survey is analysed for any signs of increased corporeal presence in this empirical study.

4.4 Part I: Results Regarding Intepretability of the Avatar's Gaze Direction

The data gathered from all thirty-one of the participants shows that the mean accuracy rate at interpreting the avatar's gaze during the control condition was 41%

with a standard deviation of 28%, while the TSB condition had a mean of 67% with a standard deviation of 20%. The standard deviations are quite high and this is most likely due to the large variability between the participants' abilities, in addition to the results containing some outliers across both experimental conditions. The outliers occurred due to the fact two participants performed worse with the TSB technique switched on and their results can be seen marked in red in Appendix A.2. Figure 4.3 shows a box plot diagram of the average accuracy rating achieved for each participant across both conditions and in Appendix A.2 the complete set of results for this experiment for each of the thirty-one participants are detailed. These results indicate that there is a higher accuracy rate being achieved by participants when the TSB technique is being used.

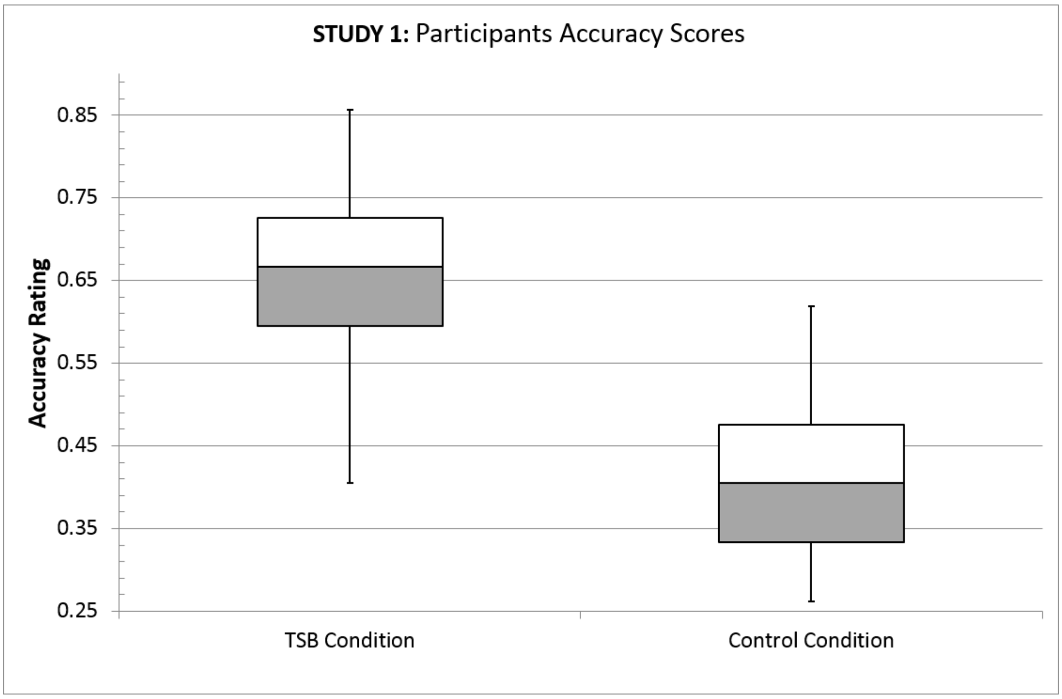


Figure 4.3: Box plot diagram highlights a portion (0.25 to 0.85) of the scale (0.0 to 1.0) for clarity and shows the average accuracy rating achieved by the participants across both conditions (i.e. *TSB* and *Control*).

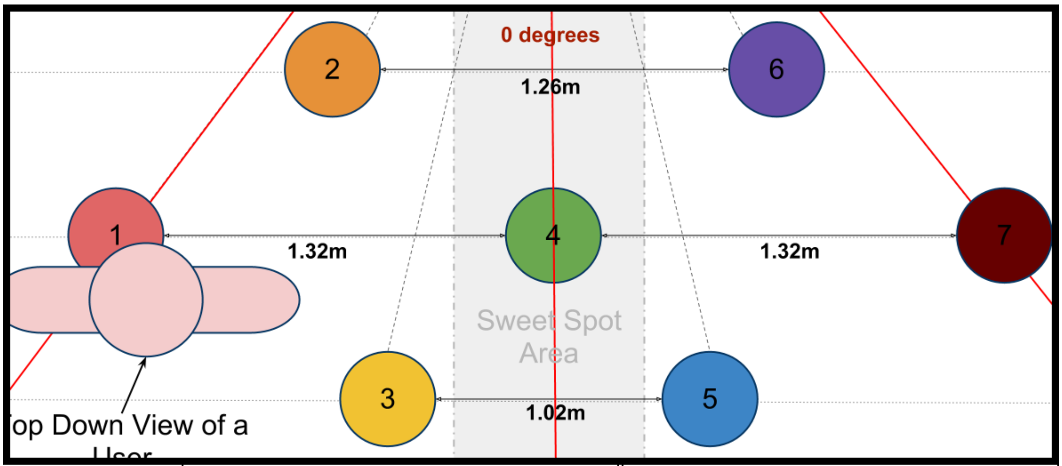
Using a *paired two sample for means Student t-Test* on the above results shows that there is a significant difference in the rating for the TSB condition (M = 0.67, SD = 0.2) and control condition (M = 0.41, SD = 0.28): $t(41) = 2.02$, $p < 0.001$. These results suggest that during the TSB condition participants tend to move to the correct marker more often than during the control condition. The higher accuracy is also an indication of an increase in Pl, meaning the participants perceived the avatar was corresponding geometrically correct to their real-world surroundings. Which further indicates that transportation, a conceptualisation of presence put forth by Lombard & Ditton (1997), holds true.

Although the results of the *paired two sample for means Student t-Test* was encouraging this analysis did not take into account the variance in the forty-two trials carried out over the seven floor markers in the experiment. In order to investigate this, an *ANOVA Two Factor with Replication* statistical test was carried out to see if there was any significant interaction occurring between the seven floor markers.

First, the control condition data from all thirty-one participants across all forty-two trials (seven samples with six trials for each floor marker) was analysed. The results show that the seven floor markers have a highly significant difference [$F(6, 1085) = 10.44$, $p < 0.001$]. This indicates a diverse set of results was gathered from the thirty-one participants across each of the seven floor markers. As expected, when the results for each participant for all forty-two trials were compared, there was also a significant difference present [$F(30, 1085) = 1.64$, $p = 0.02$]. However, there was no significant interaction between the seven floor markers and the thirty-one participants ($F(180, 1085) = 0.95$, $p = 0.68$), indicating that nothing out of the ordinary did occur during the control condition results and the results are varying greatly.

When the same *ANOVA Two Factor with Replication* statistical test was carried out on the TSB condition data gathered from the same thirty-one participants across all forty-two trials, the results once again showed a significant difference between the seven floor markers ([$F(6, 1085) = 22.48$, $p < 0.001$]) and a significant difference between all thirty-one participants ([$F(30, 1085) = 2.55$, $p < 0.001$]) which was expected. However, unlike the control condition, the TSB condition had a significant difference for the interaction between the seven floor markers and the thirty-one participants [$F(180, 1085) = 1.25$, $p = 0.02$]. This significant difference in the interaction between the floor markers and participants further indicates that a pattern was emerging, and it would seem that when the TSB technique is switched on, participants get more trials correct as they move away from the sweet spot (i.e. floor marker *4*), and a less varied and more stable set of results is observed. This pattern can clearly be seen in Figure 4.4 where the *control condition matrix diagram* shows far less accuracy when participants are moving towards floor markers on the extremities (floor markers *1* and *7*) and the accuracy improves the closer the participants get to the sweet spot (i.e. floor marker *4*). Whereas with the *TSB condition matrix diagram* the obvious stabilisation of the results can be seen across all seven floor markers further indicating the positive effect the TSB technique has on the participants' abilities to accurately interpret the gaze direction of the avatar.

The heat maps in Figure 4.4 illustrate the performance (i.e. the mean accuracy ratings for all participants) of participants at interpreting the avatar's gaze direction towards the floors markers during each trial for both experimental conditions. There are obviously more mid to high range rating blocks in the TSB matrix, indicating that the avatar achieved a greater level of geometric correspondence from the participants' perspectives in their real-world surroundings during the TSB condition. This backs up the first claim of the TSB technique as outlined in Section 1.2.



	Control Condition Matrix Diagram							TSB Condition Matrix Diagram						
	Floor Markers (moving to)							Floor Markers (moving to)						
Floor Markers (moving from)	1	2	3	4	5	6	7	1	2	3	4	5	6	7
1		0.90	0.10	0.28	0.07	0.24	0.03		0.97	0.34	0.83	0.66	0.66	0.66
2	1.00		0.21	0.45	0.21	0.45	0.10	0.90		0.31	0.83	0.41	0.83	0.66
3	0.76	0.34		0.48	0.59	0.48	0.10	0.90	0.86		0.90	0.52	0.76	0.59
4	0.59	0.48	0.66		0.48	0.62	0.45	0.41	0.90	0.62		0.45	0.90	0.28
5	0.14	0.55	0.59	0.69		0.48	0.72	0.59	0.83	0.48	1.00		0.86	0.55
6	0.24	0.55	0.10	0.41	0.07		1.00	0.72	0.76	0.48	0.72	0.52		0.79
7	0.10	0.24	0.00	0.07	0.21	0.93		0.76	0.55	0.66	0.41	0.31	0.86	

Figure 4.4: **Control condition matrix diagram**: Showing the mean accuracy rating for each of the forty-two trials gathered from all thirty-one participants during the control condition. The columns (numbered *1* to *7*) represent the floor marker the avatar was directing its gaze towards, i.e. where the participant was aiming to *move to* and the rows represent the floor markers (i.e. numbered *1* to *7*) the participants were standing on while guessing which floor marker the avatar was directing its gaze towards, i.e. where the participant was going to be *moving from*. **TSB condition matrix diagram**: Showing the mean accuracy rating for each of the forty-two trials for all thirty-one participants during the TSB condition.

The lower accuracy rating in the *control condition matrix diagram* shows that the Mona Lisa Effect does lead participants to misjudge the avatar's gaze due to a false sense of accuracy the user experiences from the direct eye contact with the avatar, up until the avatar changes its eye gaze position towards a POI. The participants accuracy at interpreting the avatar's gaze direction deteriorates the further away from the sweet spot a participant is standing while attempting to interpret the gaze direction, these findings verify the findings of the previously discussed research in Section 2.3. These results also indicate the limitations of a standard 2D display as a means to display interactive avatar interfaces, as they limit the user mobility in front of the display to the sweet spot and limit the user's ability to interpret any gaze directions from the avatar's virtual world into their own real-world surroundings.

Interestingly, participants had more difficulty interpreting the avatar's gaze when directed towards a floor marker behind where the participant was currently standing. The results also indicate that participants moving to floor markers *3* or *5*, during both the experimental conditions, achieved a lower than expected accuracy rating (i.e. floor marker *3* average percentage (TSB vs. Control): 48% vs. 28%, floor marker *5* average percentages (TSB vs. Control): 48% vs. 27%). This could be the result of floor markers *3* and *5* being the back floor markers, and nearly always being behind the participants (see Figure 4.1) when the participants had to determine the avatar's gaze direction towards these floor markers. These low mean accuracy ratings for floor markers *3* and *5* indicate difficulty for participants interpreting the avatar's gaze when the floor marker appears behind the participant.

There was an observably high accuracy rate for the participants in the control condition for moves *2 to 1* and *6 to 7*, this can be seen in Figure 4.4. This was not surprising as in the control condition participants were typically able to make broad interpretations of whether or not the avatar was looking to the *left* or *right*. When a participant was on marker *2* or marker *6* and the avatar looked right or left respectively, the participant had an easy choice to make – illustrated in the accuracy levels of 100% for moves *2 to 1* and *6 to 7*.

However, the results in Figure 4.4 also show that when the participants were on floor markers *1* or *7* for the control condition and the avatar looked to the left or right respectively, participants had a difficult choice. With the exception of the *1 to 2* and *7 to 6* moves accuracies for moves from marker *1* and marker *7* are extremely low. The high accuracies for moves *1 to 2* and *7 to 6* may have been because these were seen as the best *damage limitation* moves from marker *1* and marker *7* respectively and so were chosen to a large extent.

The results show that there is no significant difference between the control condition and the TSB condition for moves starting from floor marker *4* (i.e. the sweet spot), which can be seen in Figure 4.1 to see the position of floor marker *4* at 90° from the centre of the 2D display. A *paired two sample for means Student t-Test* shows this in the rating from floor marker *4* for the TSB condition (M = 0.59, SD = 0.26) and control condition (M = 0.54, SD = 0.09): $t(5) = 2.57$, $p = 0.66$. However, on the contrary when participants had to move from all the other floor markers which lie outside of the sweet spot (i.e. *1*, *2*, *3*, *5*, *6* and *7*), results from a *paired two sample for means Student t-Test* show a significant difference in the ratings for the TSB condition (M = 0.68, SD = 0.19) and control condition (M = 0.39, SD = 0.29): $t(35) = 2.03$, $p < 0.001$. The use of the TSB technique does seem to go a long way to compensate for the reliance on the participant to be in the sweet spot when interactive avatar interfaces are displayed on a 2D display.

The result of a *paired two sample for means Student t-Test* shows that there seems to be no statistical difference between the left-hand side moves from the sweet spot (M = 0.57, SD = 0.09), i.e. *4 to 1*, *4 to 2* and *4 to 3*, and those opposite to the right-hand side (M = 0.52, SD = 0.09), i.e. *4 to 7*, *4 to 6* and *4 to 5* for the control condition: t(2) = 4.3, p = 0.62 (see Appendix A.4 for a complete table on the symmetrical differences). The TSB condition shows similar results with the left-hand side moves (M = 0.64, SD = 0.24) and right-hand side moves (M = 0.54, SD = 0.32) showing no statistical difference: t(2) = 4.3, p = 0.19.

Similarly, a *paired two sample for means Student t-Test* shows that the left-hand side floor markers (M = 0.67, SD = 0.21), i.e. *1*, *2* and *3*, and those opposite on the right-hand side (M = 0.64, SD = 0.16), i.e. *7*, *6* and *5*, for the TSB condition have no statistical difference: t(14) = 2.14, p = 0.33. This result is repeated for the control condition for the left-hand side floor markers (M = 0.39, SD = 0.32) and those opposite on the right-hand side (M = 0.38, SD = 0.32) with no statistical difference: t(14) = 2.14, p = 0.69. These results suggest that any future results gathered in a similar experiment would be symmetrical across the right and left sides and this should impact the procedure of any further experiments.

Part I of *Study 1* has outlined the results for the gaze perception component of this experiment and the results indicate that the participants had increased accuracy at interpreting the avatar's gaze direction during the TSB condition. This indicates a higher level of Pl and thus, geometric correspondence a contributing factor of corporeal presence. This increase can also be connected with transportation, a conceptualisation of presence put forth by Lombard & Ditton (1997). The next section (see Section 4.5) details the results of the survey carried out in *Part II* of the experiment. The survey was analysed in relation to Psi, which corresponds to representation the second contributing factor of corporeal presence according to Holz *et al.* (2011).

4.5 Part II: Survey Results Regarding Perceived Corporeal Presence for the Avatar

When the survey results for both experimental conditions from Figure 4.5 are compared, the results of the comparison indicated that on average participants gave higher ratings for all the questions after completing the TSB condition. In order to test for a significant difference between both sets of survey results for each of the experimental conditions, a *paired two sample for means Student t-Test* was used. This statistical test shows that there is a significant difference in the survey question ratings for the TSB condition (M = 0.75, SD = 0.02) and control condition (M = 0.60, SD = 0.08): $t(5) = 2.57$, p[1] <0.001). The details of the questions and some explanations for participants' responses are given below:

1. *To what degree did you become so involved in doing the task that you lost all track of time?* The higher results for the TSB condition could be put down to the fact people got more moves correct and they did not have to be repositioned as often. Hence, they were more engrossed for longer periods of time throughout the experiment.

2. *To what degree did you feel the 3D virtual character's head movements were natural?* The results suggest that participants perceived that the avatar's head

[1] $p = 0.34 \times 10^{-2}$

Participant Information				Q. 1		Q. 2		Q. 3		Q. 4		Q. 5		Q. 6	
Number	Age	Sex	1st	A	B	A	B	A	B	A	B	A	B	A	B
1	*27*	*m*	B	3	4	4	5	2	5	5	5	4	4	2	2
2	*26*	*m*	A	4	5	3	5	3	4	4	4	4	4	3	4
3	*29*	*m*	B	4	4	3	3	2	3	4	4	2	4	2	3
4	*28*	*m*	A	2	3	3	3	3	4	4	3	2	4	3	4
5	*64*	*m*	B	4	5	3	4	2	3	1	1	3	4	3	4
6	*33*	*m*	A	4	4	3	4	2	4	5	3	2	2	2	3
7	*26*	*m*	B	2	4	2	3	2	4	2	5	1	4	2	3
8	*33*	*m*	A	5	3	4	4	2	3	3	2	1	3	2	2
9	*23*	*m*	B	5	3	3	5	4	4	5	5	5	3	3	5
10	*30*	*m*	A	3	3	4	5	4	4	4	4	3	4	3	4
11	*24*	*f*	B	4	2	3	3	2	3	3	5	2	3	2	3
12	*25*	*m*	A	2	3	3	4	1	4	2	2	1	3	1	4
13	*26*	*f*	B	3	4	2	5	2	3	3	5	2	4	1	3
14	*57*	*m*	A	3	1	3	3	3	4	2	2	2	2	3	4
15	*57*	*f*	B	4	5	4	4	3	3	2	5	2	4	3	4
16	*33*	*m*	A	4	4	5	5	4	5	5	4	3	4	3	4
17	*29*	*m*	B	2	3	2	4	1	4	2	3	1	2	1	1
18	*25*	*f*	A	4	5	3	4	4	4	3	4	2	3	3	4
19	*27*	*f*	B	5	5	2	3	2	3	3	3	2	3	2	3
20	*39*	*m*	A	3	4	4	5	4	5	1	4	1	4	3	4
21	*39*	*m*	B	4	4	4	4	5	5	5	5	4	5	4	4
22	*10*	*f*	A	4	5	3	4	1	5	4	3	1	4	1	5
23	*6*	*m*	B	3	4	1	3	1	3	4	4	1	3	1	3
24	*35*	*f*	A	5	5	3	3	3	3	4	4	3	3	5	3
25	*61*	*f*	B	3	3	3	4	2	4	1	5	3	4	2	4
26	*23*	*m*	A	2	2	4	2	3	4	5	4	4	4	4	4
27	*24*	*m*	B	5	5	3	4	4	2	4	4	3	4	5	5
28	*32*	*m*	A	3	4	4	4	3	4	1	4	3	4	4	4
29	*27*	*f*	B	4	3	5	4	4	4	4	5	5	4	4	4
30	*25*	*m*	A	4	3	5	4	3	3	3	4	3	3	4	3
31	*26*	*f*	B	4	4	3	5	3	3	5	4	4	4	3	4

Figure 4.5: From left to right, the first four columns shows data relating to each of the thirty-one participants: Participant **Number**, **Age**, **Sex** and the experimental condition they did first (i.e. '1^{st}' column). **A** is the *control condition* and **B** is the *TSB condition.* Then there are six columns that contain the ratings for each question (i.e. one to six) for each participant on a Likert scale of *1* to *5*. These columns are split into two, one for each of the two experimental conditions. There is a heat map effect applied where red squares represent the lowest possible rating and the green squares the highest possible rating.

movements were more natural during the TSB condition. This indicates that the Psi factor for the TSB condition would seem to be higher.

3. *To what degree did you feel the 3D virtual character's gaze direction towards the spots on the ground was realistic?* The results here are in favour of the TSB condition. This can be put down to the fact that participants had a higher accuracy during the TSB condition so they rated the avatar's gaze direction higher to reflect their own performance.

4. *To what degree did you feel the 3D virtual character was responsive to your actions?* It was predicted and the results show that there is little difference between the conditions as the avatar's responsiveness is identical for both.

5. *To what degree did your experience with the 3D virtual character's gaze seem consistent with your real-world experiences?* The difference between results for both conditions was substantial here, indicating that for the TSB condition participants on average believed that the avatar's gaze seemed more consistent with real-world experiences.

6. *To what degree do you think the 3D virtual character was actually able to look out at the real-world 'spots' on the ground?* Relating directly to the Psi factor, a higher mean accuracy rating by participants for the TSB condition indicates a higher sense of perceived corporeal presence for the avatar.

4.6 Part III: Outliers

There were two outliers in this first empirical study, participants "*8*" and "*21*" (see results highlighted with red in Appendix A.2), they both achieved a negative improvement when the TSB technique was engaged. Participant "*27*" actually scored the lowest rating for the TSB condition but that rating was a significant improvement on their control condition rating.

Participants "*8*" and "*21*" both wore glasses, were of average height and did not spend more then the average time required to complete both experimental conditions:

- Participant "*21*" was very enthusiastic throughout both experimental conditions, chatting continuously with the experimenter about "*3D illusions*" and the psychology behind them. It is quite possible that this behaviour may of been a factor in their low scores across both conditions. However, participant "*21*" seems to have been unaware of their poor performance and rated both experimental conditions equally high for the survey questions (see survey results highlighted with red in Appendix A.5 and A.5), which is hard to explain.

- Participant "*8*" on the other hand was very disinterested in taking part in the second experimental condition (i.e. TSB condition for them). This was most likely due to their perceived poor performance during the control condition (i.e. their first experimental condition) previously, which was evidently the cause of frustration for them. During the TSB condition participant "*8*" second guessed themselves continuously, often changing their mind from the correct floor marker to an incorrect one at the last minute. This ultimately lead to more frustration for this participant and can be seen in their poor ratings for the survey questions (see results highlighted with red in Appendix A.5 and A.5).

It was not deemed necessary to remove participants "*8*" and "*21*" from the analysis as they both carried out the experimental conditions with no problem, they just may not of been ideal participants and were not invited back for any further experiments.

4.7 Discussion

The results of *Study 1* highlights two important observations. Firstly, in support of the first claim of the TSB technique (see Section 1.2), participants achieved a higher rate of accuracy at interpreting the avatar's gaze direction across all seven floor markers during the TSB condition with no obvious degradation of accuracy the further away from the sweet spot participants stood. This is a further indication that Pl was experienced by participants meaning that the avatar was corresponding geometrically correct to its real-world surroundings from the participants' perspectives. Secondly, in support of the second claim of the TSB technique (see Section 1.2), the survey results indicate some significant difference between the ratings for the questions across both conditions. On average participants rated the survey questions higher after the TSB condition, regardless of the fact that the avatar performed exactly the same behaviour throughout both experimental conditions. This indicates an increase in Psi (Slater, 2009) and subsequently, representation (Holz *et al.*, 2011). This increase in Psi in addition to the increase of Pl signifies an overall increase in the perceived corporeal presence for the avatar. Lastly, it was noted that using a standard 2D display for an avatar UI limits the user's ability to fully interpret gaze direction and hence, limits the avatar from being able to accurately direct the user's visual attention to PoI in the user's real-world surroundings. However, with the addition of the TSB technique this would no longer seem to be the case and as such all three of these observations can be related to the contribution of this thesis as outlined in Section 1.2.

Another observation made during *Study 1* paves the way for empirical *Study 3* (see Chapter 5), where the question of using 3D display technology in conjunction with a 3D avatar UI would have any bearing on a participant's ability to interpret the avatar's gaze direction. As 3D display technology will only render the image to stereoscopic, participants will still be required to be in the sweet spot to view the rendered 3D image from the correct perspective in order for the 3D illusion to hold true. Hence, having no effect on a participant's ability to interpret the avatar's gaze direction from outside of the sweet spot. A graphical process such as the TSB technique is still required in order to combat lateral foreshortening. In Chapter 6 the two additional experimental conditions are outlined, *3D Control* and *3D TSB*, as they were used to evaluate if 3D display technology had any bearing on the participant's ability to interpret gaze direction in conjunction with the TSB technique or just by itself.

However, the results from *Part I* of *Study 1* suggest that the TSB technique is a good first step in enabling an avatar to deliver accurate gaze direction from the 2D display into the participant's real-world surroundings from the participant's perspective. Furthermore, this implies there is a high level geometric correspondence between the avatar and the participant's real-world surroundings. A high level of geometric correspondence is important as it contributes to 50% of the perceived corporeal presence of an avatar. *Study 2* in the next chapter (see Chapter 5) examines this aspect of the TSB technique and further evaluates the first claim of the TSB technique (see Section 1.2) with a more rigorous experimental set-up.

Chapter 5

Empirical Study 2: Find the '*Sweet Spot*'

The results from *Study 1* (see Chapter 4) show that during the TSB condition a participant can interpret the direction of the avatar's gaze direction more accurately than they can during the control condition. However, the results from *Study 1* also show that there was no significant difference between the control condition and the TSB condition when the participant is standing in the sweet spot (i.e. floor marker *4* located at optimal 90° viewing angle to the centre of the 2D display in *Study 1*, see Figure 4.1).

The results of *Study 1* show that there is no significant difference between the left-hand side (see floor markers *1*, *2* and *3* in Figure 4.1) and right-hand side (see floor markers *5*, *6* and *7* in Figure 4.1) floor markers (excluding floor marker *4* in the sweet spot). Also, the results from *Study 1* indicate that participants had difficulty at interpreting the floor markers (i.e. specifically floor markers *3* and *5* from *Study 1*) the avatar was directing its gaze towards when the floor markers were physically behind the participant who was trying to guess the avatar's gaze direction towards them. Taking all the results from *Part I* of *Study 1* regarding the participant's increased accuracy at interpreting the avatar's gaze direction, the design of this follow up study would need to investigate how much more the TSB technique could improve a participant's ability to interpret the gaze direction of the avatar across a larger area within the participant's real-world surroundings. The number of PoI increased, while their placement in the participant's real-world surroundings exceeded the safe area (i.e. the Kinect's FoV at about 57°) seen in *Study 1*, which contained the seven floor markers. This expansion of the number and spread of the PoI throughout the participant's real-world surroundings will help to verify the findings from *Study 1* by further evaluating the first claim of the TSB technique in a more elaborate experimental set-up (see Section 1.2). This empirical study is evaluating the TSB techniques ability to increase the user's ability to use gaze perception to interpret the avatar's gaze direction towards PoI in the participant's immediate surroundings.

The increase in interpretability can be put down to an increase in Pl which means the avatar is achieving a greater geometric correspondence with the participant's real-world surroundings from the participant's perspective. This is what makes up 50% of corporeal presence according to Holz *et al.* (2011). Furthermore, as the participant's sense that the avatar was geometrically corresponding correctly to its immediate surroundings increased, the greater the likelihood that the conceptualisation of presence as transportation (Lombard & Ditton, 1997) will also increase.

Where *Study 2* differs from *Study 1* is that the participants had many more PoI in their immediate surroundings in which the avatar could direct its gaze towards. In *Study 1* there were only seven floor markers which were all placed within the FoV of the Microsoft Kinect, referred to in this study as the *free move area* which can be seen in Figure 4.1. The free move area is marked by white tape in Figure 5.2 and is outlined by a black line in Figure 5.1. In this empirical study the free move area is the space in which the participant can freely move within while they attempt to interpret the avatar's gaze direction, which allows the user to move naturally while trying to interpret the avatar's gaze direction. This is in contrast to *Study 1* where the participant had to remain stationary on the floor marker they were standing on at the start of a trial while they interpreted the avatar's gaze direction, this empirical study removed these limitations.

In this empirical study the participant had to alternate between two starting positions, a red square and a green square which are illustrated in Figure 5.1 and were marked out on the floor in the laboratory by red and green tape as seen in Figure 5.3. These starting positions ensured that the participant started each trial (i.e. with the participant standing in either the green or red square starting positions and stated they were ready, then the avatar would direct its gaze towards a PoI) from outside of the sweet spot. However, once the avatar had directed its gaze towards a PoI the participant was free to move anywhere within the free move area before making their guess as to which PoI the avatar was looking at. The decision was made to allow the participant to freely move around as it was a more natural process than standing stationary as participants did in *Study 1*, while interpreting the avatar's gaze direction. There was also the addition of multiple distraction PoI, placed in the participant's vicinity in this empirical study to ensure it was not a trivial task. *Study 1* had only seven PoI which can be seen in Figure 4.1, whereas this empirical study had thirty PoI.

The thirty PoI were also split into four different groups and colour-coded accordingly, as the colour coding makes it easier to reference PoI, to design the procedure of the experiment and to organise the gathered data. The colour-coded groups can be seen in Figure 5.1. Participants were not aware of these groups or their colours. Only the eight wall mounted PoI were visible with labels at all times during the experiment. The four groups are as follows:

1. **Black**: These eight PoI were all placed on the walls of the laboratory at eye level and were labelled with letters, *A* to *H* as seen in Figure 5.1. The placement of these PoI was to simulate distant objects, such as doors and windows within the participant's vicinity.

2. **Purple**: This PoI group was made up of eight floor PoI that are placed close to the 2D display on which the avatar appears and are placed outside of the participant's free move area to the far right and left. This group helped to determine how well participants can interpret a PoI placed between the avatar and themselves, as well as the user ability to interpret the avatar gaze direction as it looks to the far right and left of the participant's current position.

3. **Red**: This PoI group was made up of seven floor PoI that are placed in exactly the same locations as the original seven floor markers from *Study 1* seen in Figure 4.1. This group were situated in the participant's free move area which

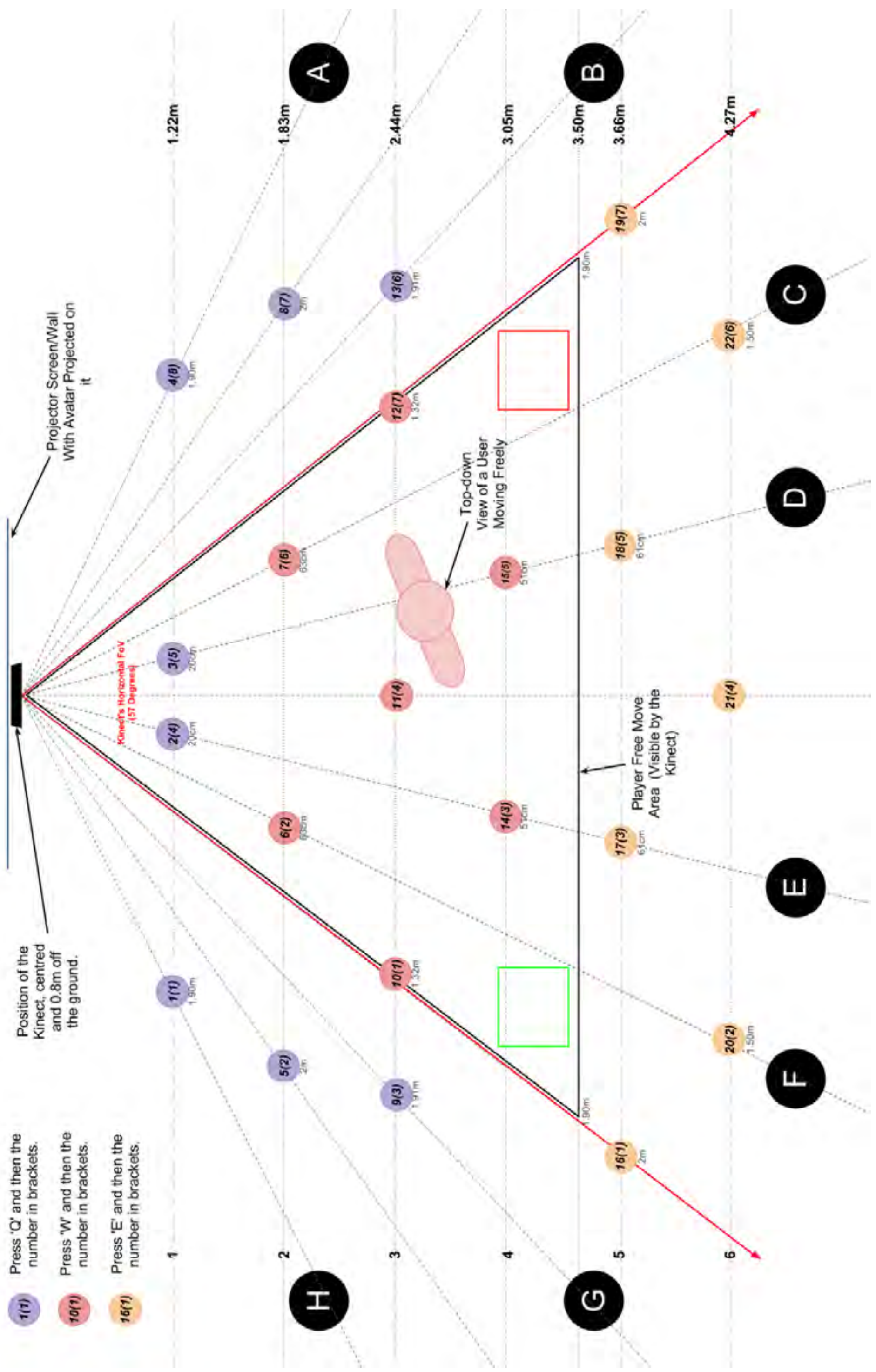


Figure 5.1: This diagram illustrate a top-down view of the layout of the PoI in *Study 2*. The twenty-two floor markers (colour-coded into three groups: *Purple*, *Red* and *Yellow*), the eight wall mounted PoI (colour-coded *Black*), the free move area represented by the black triangle and the two starting positions (the green and red squares within the free move area). The FoV of the Kinect sensor (57°) is represented by the red lines. There is information in the top left corner for the experimenter who controlled this *Wizard of Oz* experiment (Dahlbäck *et al.*, 1993).

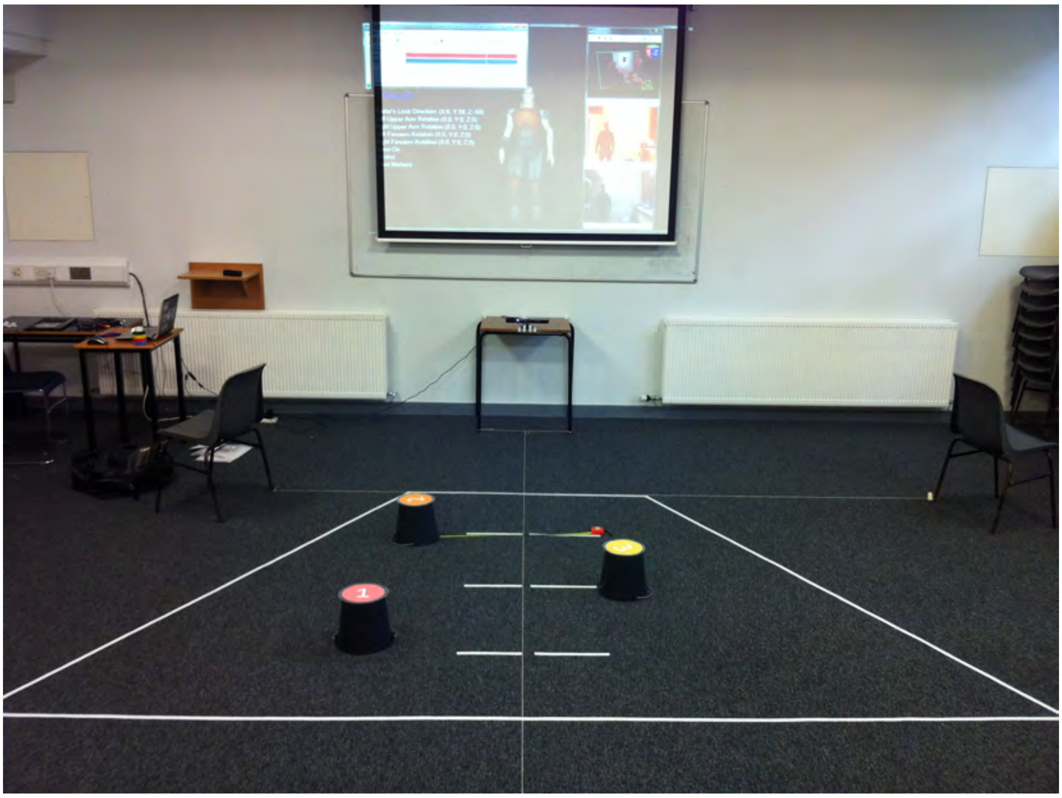


Figure 5.2: Measuring out and marking the floor according to Figure 5.1.

allowed the participant freedom to move around the free move area containing the PoI while interpreting the avatar's gaze direction towards one of the PoI.

4. **Yellow**: These seven PoI were placed behind the participant's free move area in order to see if the participant can interpret PoI that are behind them as they look at the avatar in order to determine its gaze direction. *Study 1* results indicated a higher rate of inaccuracy occurred when participants were guessing the floor makers that appeared behind them. The results show low accuracy for moves to floor markers *3* and *5* (i.e. the furthest back floor markers in *Study 1*) and the results can be seen in Figure 4.4.

Each of the PoI on the floor that made up the *Purple*, *Red* and *Yellow* PoI groups (see Figure 5.1) were marked with blue stickers. In Figure 5.2 the laboratory is being measured up and marked out to ensure the PoI are in the correct positions for the experiment. However, there were additional blue stickers inter-mingled throughout the floor area amongst the PoI. These were to form part of a distraction tactic employed to ensure that the participants were prevented from seeing any obvious pattern on the floor and deciphering the arrangement of pre-defined PoI on the floor. Only the experimenter knew which blue stickers related to the PoI (i.e. the pre-defined PoI from the *Purple*, *Red* and *Yellow* groups) that were scripted for each path to contain a hidden prize (i.e. sweets/candy) during each of the trials and covered with one of the three upturned buckets.

5.0.1 The Two PoI Types

The two types of PoI in this empirical study are wall mounted PoI at eye level to the participants and floor based PoI similar to *Study 1* but unmarked and greater in number. These two distinct PoI types were evaluated using two categories of trials, which are as follows:

1. *The Bucket Guessing (BG) Trials*: Associated with all the floor based PoI in the *Purple*, *Red* and *Yellow* groups. Figure 5.3 shows two participants guessing which upturned bucket the prize is under. The first participant in **1A** starts the trial from the red square starting position (see the red and green square starting positions in Figure 5.1). In **1B** the avatar has directed its gaze towards a PoI (one of three buckets labelled *1*, *2* or *3*), an upturned bucket on the floor covered the hidden prize (i.e. sweets/candy). Furthermore, the two neighbouring upturned buckets not hiding a prize were purely distractions, in

order to make the experiment more challenging while testing the accuracy of the participants' abilities to interpret the avatar's gaze direction. Once the avatar has directed its gaze towards the PoI the participant is free to move within the free move area in order to make their guess as to which upturned bucket conceals the prize. In **1C** the participant has made their choice of bucket and indicates their choice vocally to the experimenter before they are free to pick up the bucket to reveal if their guess is correct, which means they win a prize. Then the participant has to go to an adjoining room and wait for the experimenter to reset the buckets for their next trial.

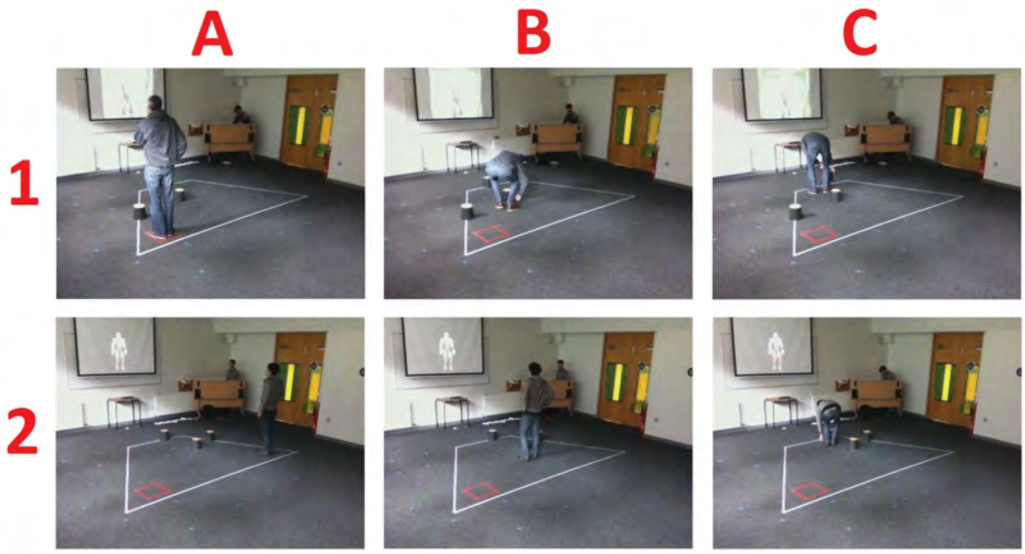


Figure 5.3: This figure shows two participants guessing which bucket the avatar is directing its gaze towards in order for the participant to guess where the hidden prize is. In row '**1**' the participant is carrying out a trial with the TSB condition switched on and the row '**2**' participant is doing a trial with the control condition switched on. In column '**A**' both participants are waiting in the green or red square starting position for the avatar to initiate the trial by directing its gaze towards one of the buckets. Column '**B**' shows the participants engaged in guessing which bucket the prize might be under, once the avatar directs its gaze the participants are free to move within the free move area. Lastly, column '**C**' shows the participants guessing which bucket the prize is under, they call out the number of the bucket (i.e. *1*, *2* or *3*) then they proceed to lift that bucket. If they guessed correctly they get the prize and then the room is reset for another trial while the participant waits in an adjoining area where they can't see the experimenter placing the prize beneath a bucket.

2. *The Letter Guessing (LG) Trials*: Associated with all wall mounted PoI in the *Black* group only. This part of the experiment is similar to the set-up used by Gibson & Pick (1963) where the participant had to guess where the actors gaze had been directed behind them. Once a participant had finished a set of four BG trials they had to do one or two LG trials to the predetermined *Path* (see Appendix B.1 and B.2 data sheets). The participant would be informed by the experimenter that they had to do an LG trial and that they were to remain in the free move area (see Figure 5.1). When the participant acknowledged they were ready, the avatar would direct its gaze towards a letter on the wall. All eight letters were looked at once during the experiment with no repeats. The participant would tell the experimenter what letter they thought the avatar was looking at and the experimenter would take note of their answer. This was repeated a second time or the participant would go back to the adjoining screened off waiting area while the experimenter set up the next BG trial. The LG trials helped break up the task of doing the BG trials and was a secondary task within the experimental process.

5.0.2 The Bucket Placements

A *compass*-styled bucket placement system based on four 90° angles with three positions to place the buckets was devised to ensure that each participant got the same bucket placement for each trial, achieving consistency in gathered data. Figure 5.4 depicts all of the possible bucket placements in the '*compass*-styled' bucket placement system. The four 90° angles were as follows:

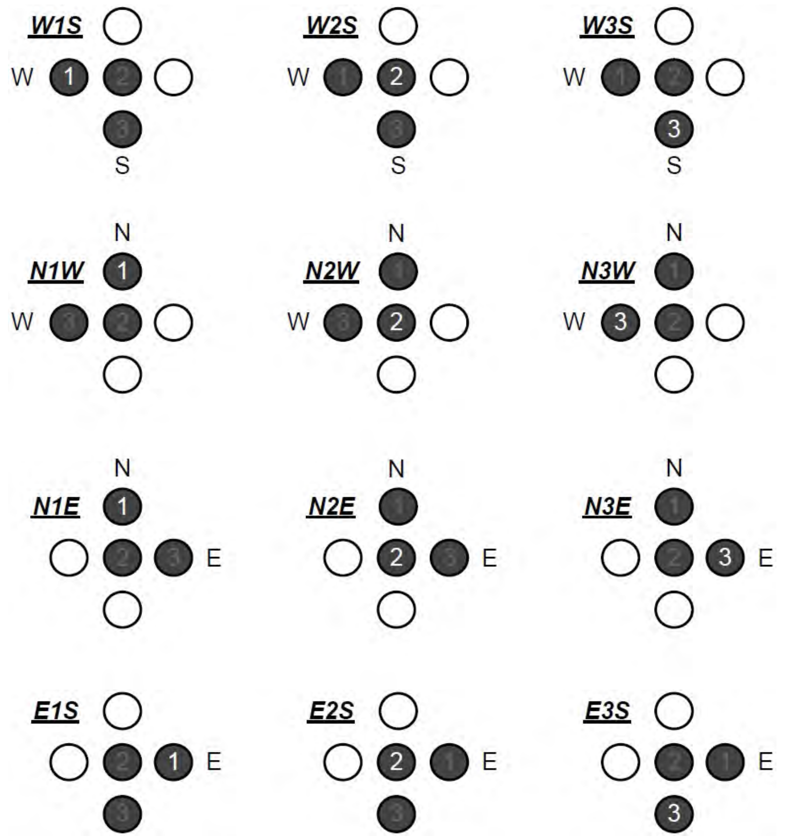


Figure 5.4: This compass styled bucket placement scheme was based on right angles, for example *W1S* has the number *1* highlighted in white and it is in the *West* position. This meant the hidden prize was to be placed under this bucket, i.e. the PoI being examined has the hidden prize on it and was to have two empty buckets in position *2* which is always the centre of the compass and *3* the *South*. *South* would always be the furthest bucket placement from the 2D display.

- **West-South**: This 90° angle encompasses three points, they were *West*, *Centre of the Compass* and *South*. The hidden object could appear on one of these three points and trials were randomly assigned an order.

- **North-West**: This 90° angle encompasses three points, they were *North*, *Centre of the Compass* and *West*. Again, the hidden object could appear on one of these three points and trials were randomly assigned an order.

- **East-North**: This 90° angle encompasses three points, they were *East*, *Centre of the Compass* and *North*. Again, the hidden object could appear on one of these three points and trials were randomly assigned an order.

- **East-South**: This 90° angle encompasses three points, they were *East*, *Centre of the Compass* and *South*. Again, the hidden object could appear on one of these three points and trials were randomly assigned an order.

This system created twelve unique bucket placement arrangements and then one of the twelve was assigned to each of the trials (see the column labelled *Placement* in Appendix B.1 *Path 1* and Appendix B.2 *Path 2*). This ensured that each bucket was always placed on the same blue stickers for each of the trials for all of the participants.

5.1 Procedure

Similar to *Study 1* (see Chapter 4), the second empirical study was based on *Wizard of Oz* experimental set-up (Dahlbäck *et al.*, 1993). Ensuring that each participant had a similar and consistent experience to every other participant throughout their trials and the avatar's gaze behaviours were again controlled by the experimenter. Each participant had to sign a release form (see Appendix B.3) before they were issued with instructions to carry out the thirty-two trials based on the thirty PoI in the *Purple*, *Red*, *Yellow* and *Black* groups. The extra two trials are due to the placement of two PoI (one in the *Red* group and one in the *Yellow* group) in the sweet spot. This meant each participant had to do each of these PoI twice, once for both the TSB and control condition (the experimental conditions were identical to those of *Study 1* outlined in Chapter 4). The issued instructions were as follows:

1. You will be required to sit in the holding area (*experimenter note*: point to screened off area in the corner of the room), where you will not be able to see the experimental area, when I (the *experimenter*) say ready you come out of the holding area and stand on either the green or red square starting position. I will specify which square and they will alternate throughout the experiment.

2. When you are in a starting position the avatar will appear on the screen and you will have to verbally say you are "*ready*" before the avatar will direct its gaze toward one of the three upturned buckets on the floor.

3. Once the avatar has directed its gaze towards one of the buckets you are free to move within the white line marked area.

4. You can take as long as you need to guess which bucket the avatar is looking at.

5. Once you are happy with your guess, just pick up the bucket you have chosen. If there is some candy underneath the bucket you have won, if not, too bad. Either way you then have to return to the holding area until you are called for the next trial.

6. After every four of these bucket guessing trials you will be asked by me (the *experimenter*) to remain within the white line marker area. When you are ready the avatar will then look at one of the eight letter mounted on the wall.

You are free to move about the white line marker area in order to guess and when you are happy with your guess just call out the letter.

7. The letter guesses happen once or twice for every four bucket guesses and after the letter guesses you must return to the holding area to be called for the next bucket guess.

8. Feel free to ask a question or ask for a break at any time during the experiment.

9. This experiment will take less than twenty minutes but there is no time limit so go at your own pace.

10. Finally, all the prizes you win are yours to keep and take home, if you want you can eat some of the candy while you are in the holding area between trials.

The experiment would begin with a floor based trial where the participant would be standing on either the green or red square starting position and once the participant said they were "*ready*", the avatar would then direct its gaze towards a PoI on the floor. The participants would start on the opposite square for each consecutive trial, e.g. red, green, red, green, red, etc., etc. The green and red square starting position were deliberately placed outside the sweet spot to ensure participants would experience the negative effects of viewing a 2D display from an acute viewing angle. Once the avatar directed its gaze towards a PoI the participant was free to move off of their starting square and move anywhere they liked within the free move area.

In Figure 5.3 there are a series of images illustrating two participants while they guessed which bucket the hidden object was under. The buckets were labelled *1*, *2* and *3* but this had no significance on the outcome of the trial. Once the avatar was directing its gaze at the correct bucket the participant was free to move off of their starting position and move about the free move area in order to help them interpret the avatar's gaze direction. When they had selected the bucket they were free to pick it up to reveal if they won the prize or not. In either case the experimenter would ask the participant to leave the laboratory so they could set up the next trial according to a pre-defined sequence.

The experiment was structured so that a participant would carry out four BG trials for the PoI in the *Purple*, *Red* and *Yellow* groups, then one or two of the LG trials depending on the predetermined path (see Appendices for *Path 1* B.1 and *Path 2* B.2). Each of the odd numbered participants did *Path 1*, while the even numbered participants did *Path 2*. This meant that after every two participants had finished the experiment there was a complete set of data gathered for *Path 1* and *Path 2*, which meant the entire set of thirty-two PoI was covered for both conditions making this experiment '*between subject*' opposed to *Study 1* (see Chapter 4) which was '*within subject*'. This approach was decided upon after the evaluation of the results from *Study 1* showed that there was no significant difference in the participants accuracy ratings between the symmetrical PoI (i.e. in *Study 1* *2* was symmetrical to *6*, *1* to *7* and *3* to *5*) during each of the experimental conditions. So by not having every participant carry out both experimental conditions for each PoI, this meant a large saving on the duration of time the experiment took to complete for each participant.

The two experimental conditions in the empirical study mirror the conditions from *Study 1* (see Section 4.1) and are as follows:

- **Control condition**: The avatar appears as it would on a regular 2D display, i.e. the rendering does not update to reflect a participant's position.

- **TSB condition**: The TSB technique is switched on, therefore the avatar's 3D rendering is continuously updated to reflect the participant's perspective.

During each experiment every participant did a selection of floor based PoI then a wall based PoI and then back to another selection of floor based PoI again. This cycle continued until the participant covered all thirty PoI. As the thirty PoI were symmetrical and the results from *Study 1* showed no significant differences between markers on the left and right-hand sides of the screens, the participants in *Study 2* covered one half of the PoI with the control condition on and the second half with the TSB condition turned on. This meant that the experimental condition was constantly changing through the experiment and had two advantages over *Study 1*. First, the speed at which the experiment was conducted was increased, as participants had to do less trials and second, the participant had a good mix of winning (i.e. most likely during the TSB condition according to the results of *Study 1*) and losing (i.e. most likely during the control condition according to the results of *Study 1*) so their interest in carrying on the experiment remained at an enthusiastic level throughout. This different approach was inspired by *Study 1* where participants carrying out the control condition found the repetitive nature of guessing incorrectly or at least thinking they had guessed incorrectly quite frustrating. During *Study 1* some participants figured out that when a message was displayed asking them to move to a different floor marker it was because they had just moved to the incorrect floor marker. Participants having to constantly move and then move again, added time to the duration of the experiment as well as frustration, which was more apparent in participants who did the control condition second in *Study 1.* The order of which experimental condition was used with which PoI was predefined and the sequence of trials ensured 50% of the control condition trials started on the red square starting position and the other 50% started on the green square starting position, likewise for the TSB condition. The procedure used to carry out this experiment is a more elaborate than that of *Study 1*, however, they both are essentially testing each participant's ability to interpret the avatar's gaze direction. This empirical study's set-up goes further to evaluate the first claim of the TSB technique (see Section 1.2).

It is important to note that no survey was conducted as part of this empirical study, this was for two reasons: (1) the experimental conditions were interchanged throughout each experiment with every participant, so survey questions would be of no use in determining the participants experience of corporeal presence with either condition, and (2) the avatar UI was exactly the same in every way to the avatar UI from *Study 1* so it was expected that the same ratings as before with no additional insights would be retrieved, if the conditions were carried out separately.

5.2 Participants

There were fifteen participants in total (four females, eleven males) with ages ranging from twenty-five to sixty-three years. Like *Study 1* this age group was selected as it is representative of the wide range of visitors to a museum or other large public building. Each participant signed a release form (see Appendix B.3) before being issued with instructions (see Section 5.1) and eventually carrying out the experiment. Participants did not have to answer any survey questions at the end of the experiment but gave permission in the release form to allow video recording to take place while they carried out the experiment. This video footage documents how participants behaved while interpreting the avatar's gaze direction.

5.3 Results

Unlike *Study 1* with only seven PoI, this empirical study had thirty PoI split into four groups (*Purple*, *Red*, *Yellow* and *Black*). Both the LG (i.e. *Black* PoI group only) and BG (i.e. *Purple*, *Red* and *Yellow* PoI groups) trails had only one of two possible outcomes, *correct* ('*1*') or *incorrect* ('*0*'), just like *Study 1.* So all thirty PoI will be treated the same in the analysis of the results, all thirty PoI in this empirical study are just predefined points, like the seven floor markers were in *Study 1*, in the participant's vicinity that the avatar could direct its gaze towards. The avatar's behaviour was exactly the same for all the trials (BG and LG trials), again the avatar's behaviour remained unchanged from *Study 1.* The design of this empirical study was based on the findings of *Study 1* where it was noticed that there was no significant difference in the results for symmetric PoI on the left-hand side and right-hand side of the sweet spot, across each experimental condition. This meant that this empirical study is designed so that all the gathered data from the two paths participants followed (i.e. Path 1 or Path 2) could be merged to give a complete set of data for both experimental conditions for all thirty PoI and benefited participants with regard to reducing the duration of the experiment as previously mentioned.

A series of five *paired two sample for means Student t-Tests* were carried out separately on the results of the four PoI groups for both experimental conditions and are as follows:

1. This first *paired two sample for means Student t-Test* shows that a statistically significant difference exists between the *Control* condition (M = 0.5, SD = 0.23) and the *TSB* condition (M = 0.71, SD = 0.2): t(29) = 2.05, $p < 0.001$ for all thirty PoI in the study (see the box plot diagram in Figure 5.5 depicting these results).

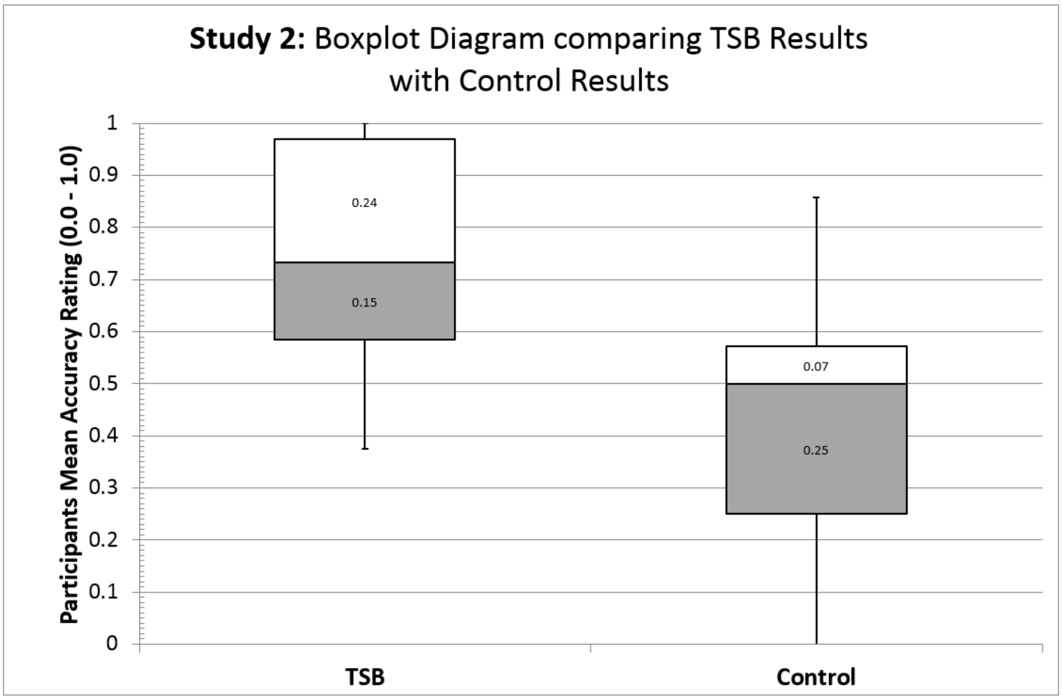


Figure 5.5: Box plot diagram showing the accuracy rating for each participant across both conditions, TSB and control.

2. **Purple PoI Group**: The results of the *paired two sample for means Student t-Test* comparing experimental conditions carried out on the data in Table 5.1 show a statistically significant difference between the control condition (M =

0.49, SD = 0.26) and TSB condition (M = 0.86, SD = 0.17): t(7) = 2.36, p < 0.01.

Table 5.1: The mean accuracy rating achieved by the fifteen participants for each of the eight PoI in the *Purple* group for the two experimental conditions. The mean accuracy rating for each condition is also shown. The column on the left, *Floor Marker*, corresponds to the floor markers for each PoI group (see Figure 5.1).

Floor Marker	**Control**	**TSB**
1(1)	0.5	0.71
5(2)	0.86	1.0
9(3)	0.0	1.0
2(4)	0.57	1.0
3(5)	0.63	0.71
13(6)	0.57	0.88
8(7)	0.25	0.57
4(8)	0.57	1.0
Mean	**0.49**	**0.86**

3. **Red PoI Group**: The results of the *paired two sample for means Student t-Test* comparing experimental conditions carried out on the data in Table 5.2 show a statistically significant difference exists between the control condition (M = 0.38, SD = 0.21) and TSB condition (M = 0.76, SD = 0.14): t(6) = 2.45, p < 0.01.

4. **Yellow PoI Group**: The results of the *paired two sample for means Student t-Test* comparing experimental conditions carried out on the data in Table 5.3 show a statistically significant difference exists between the control condition (M = 0.24, SD = 0.26) and TSB condition (M = 0.62, SD = 0.22): t(6) = 2.45, p = 0.04.

Table 5.2: The mean accuracy rating achieved by the fifteen participants for each of the seven PoI in the *Red* group for the two experimental conditions. The mean accuracy rating for each condition is also shown. The column on the left, *Floor Marker*, corresponds to the floor markers for each PoI group (see Figure 5.1) and is identical to the placement in *Study 1* (see Figure 4.1) and *Study 3* (see Figure 6.2).

Floor Marker	**Control**	**TSB**
10(1)	0.57	0.75
6(2)	0.25	0.57
14(3)	0.43	0.63
11(4)	0.6	0.79
15(5)	0.13	0.71
7(6)	0.57	1.0
12(7)	0.13	0.86
Mean	**0.38**	**0.76**

Table 5.3: The mean accuracy rating achieved by the fifteen participants for each of the seven PoI in the *Yellow* group for the two experimental conditions. The mean accuracy rating for each condition is also shown. The column on the left, *Floor Marker*, corresponds to the floor markers for each PoI group (see Figure 5.1).

Floor Marker	**Control**	**TSB**
16(1)	0.0	0.57
20(2)	0.43	0.75
17(3)	0.14	1.0
21(4)	0.14	0.38
18(5)	0.25	0.57
22(6)	0.0	0.71
19(7)	0.71	0.38
Mean	**0.24**	**0.62**

5. **Black PoI Group**: The results of the *paired two sample for means Student t-Test* comparing experimental conditions carried out on the data in Table 5.4 show that no statistically significant difference exists between the control condition (M = 0.53, SD = 0.1) and TSB condition (M = 0.76, SD = 0.22): t(7) = 2.36, p = 0.06.

Table 5.4: The mean accuracy rating achieved by the fifteen participants for each of the eight PoI in the *Black* group for the two experimental conditions. The mean accuracy rating for each condition is also shown. The column on the left, *Floor Marker*, corresponds to the floor markers for each PoI group (see Figure 5.1).

Floor Marker	**Control**	**TSB**
A	0.5	0.57
B	0.43	1.0
C	0.5	0.86
D	0.57	1.0
E	0.5	0.71
F	0.43	0.88
G	0.63	0.71
H	0.71	0.38
Mean	**0.53**	**0.76**

The results of the first *paired two sample for means Student t-Test* comparing the entire list of thirty PoI across all four groups (*Purple*, *Red*, *Yellow* and *Black*) for both the control and TSB conditions, indicate that there was a significantly higher percentage of accuracy of gaze interpretations made by participants when the TSB technique is used.

Furthermore, the results of the second, third and fourth t-Tests show significant differences between both experimental conditions for the *Purple*, *Red* and *Yellow* groups. This means for all the BG trials the TSB technique increased the participants' abilities to interpret the avatar's gaze direction towards the correct bucket on the floor.

However, the results of the fifth and final *paired two sample for means Student t-Test*, comparing the results of the two experimental conditions for the *Black* PoI group show there was no significant difference. This can be explained by the fact the process for interpreting the avatar's gaze for the *Black* PoI group was slightly easier than the other three PoI groups. As a secondary task to the BG trials, participants had to guess which one of the eight letters the avatar was directing its gaze towards. The eight letters were evenly spaced around the walls of the laboratory. It would have been more challenging for the participants if there was more letters on the wall to act as distractions when the participant was interpreting the avatar's gaze direction. Also the placing of all the letters should be less uniform to prevent predictability.

Table 5.5: The mean accuracy rating achieved by the fifteen participants for each of the six *Red* group PoI (same exact positions as the floor markers from *Study 1*) outside of the sweet spot (i.e. *11(4)*) for the two experimental conditions. The mean accuracy rating for each condition is also shown.

Floor Marker	**Control**	**TSB**
10(1)	0.57	0.75
6(2)	0.25	0.57
14(3)	0.43	0.63
15(5)	0.13	0.71
7(6)	0.57	1.0
12(7)	0.13	0.86
Mean	**0.35**	**0.75**

The results of *Study 1* (see Section 4.4) show that there was no statistically significant difference between both experimental conditions for trials that began with the participant standing on floor marker *4* (i.e. the sweet spot), while the avatar directed its gaze towards the other six floor markers. This is opposed to the statistically significant difference seen when participants started on any of the six other floor markers in *Study 1*. In this empirical study all the trials began outside the sweet spot either to the right or left in the green or red square starting positions. When the results from Table 5.5 for the six *Red* group PoI outside of the sweet spot across both conditions are compared using a *paired two sample for means Student t-Test* a statistically significant difference exists – control condition (M = 0.35, SD = 0.21) and TSB condition (M = 0.75, SD = 0.16): $t(5) = 2.57$, $p < 0.01$. This indicates that in order to gain any advantage from the sweet spot (see Section 4.4 for the mirroring results in *Study 1*) when a participant is interpreting the avatar's gaze direction, the participant must be in the sweet spot before the trial begins (i.e. before the avatar directs its gaze towards a PoI). This is evident in *Study 1* where during the control condition participants managed to score higher than average accuracy rating for trials that started on floor marker *4* (i.e. the sweet spot).

5.4 Discussion

Just like *Study 1* (see Chapter 4) the quantitative results for the twenty-two PoI placed on the floor (i.e. *Purple*, *Red* and *Yellow* PoI groups) in this empirical study are pretty definitive and show that during the TSB condition participants saw an increase in their ability to accurately interpret an avatar's gaze direction. These findings further support the first claim of the TSB technique (see Section 1.2). However, unlike *Study 1* the experimental set-up in this study had far more PoI, as well as distractions (i.e. additional blue stickers on the floor and the two upturned buckets that did not conceal the hidden prize) within the participants' surroundings to make for a more thorough investigation of the TSB technique. There is some evidence, although not statistically significant that the TSB technique led to higher accuracy for the PoI mounted on the walls in the *Black* group. This may be due to the set-up of the LG trials being relatively easy in comparison to the BG trials.

This group requires further investigation with a more elaborate experimental set-up and possibly more distraction PoI placed in the vicinity of the eight wall mounted letters (i.e. additional letters).

The results of the BG trials indicate that participants had an easier time interpreting the avatar's gaze direction when it was being directed towards PoI closer to the display, with the *Purple* PoI group achieving higher accuracy ratings than the *Red* and *Yellow* PoI groups. In turn, the *Red* PoI group achieved higher accuracy ratings than the *Yellow* PoI group and so on. The further the PoI group was from the 2D display the lower the accuracy ratings were, this can be seen in Table 5.6. This pattern was found to occur across both experimental conditions. One of the conclusions was that this pattern occurred due to the calibration of the Microsoft Kinect, its height above the ground at about one metre and the pitch angle of the Kinect pointing towards participants. From the Kinect's perspective as the participants moved further away, the actually vertical height appeared to reduce, this probably could be combated by positioning the Kinect higher off of the ground and ensuring the pitch angle is perfectly square to the participants. However, this also would mean partially obstructing the participants' view of the avatar on the 2D display with the Kinect appearing in front of the 2D display. This requires further investigation as there may be a coding fix for this issue, where a variable is added to (or subtracted from depending on the scenario) the vertical height of the participant captured by the Kinect, this variable could be calculated from a participant's distance to the Kinect and the pitch angle of the Kinect. That being said, there could also be some unknown factors at play here and only further experimentation will weed out the true cause of this observed pattern.

Table 5.6: The overall mean accuracy rating for both conditions achieved by the fifteen participants in the *Purple*, *Red* and *Yellow* groups decreased as the groups moved further away from the 2D display.

Group	**Control**	**TSB**
Purple	0.49	0.86
Red	0.38	0.76
Yellow	0.24	0.62

The experimental set-up in this study eliminated the advantage the control condition had in *Study 1*, where the results from participants starting on floor marker *4* (i.e. the sweet spot) showed higher than average accuracy ratings. The experimental set-up in this empirical study forced participants to start well outside the sweet spot in the most extreme right and left positions the Kinect's FoV (about 57°) would allow while still being able to track the participant. This ensured the participant started each trial in a position that would invoke the problem encountered with 2D display (see Section 1.1) when viewed from acute angles outside of the sweet spot. Once the avatar initiated a trial by directing its gaze towards one of the thirty PoI, the participants were then free to move anywhere they liked within the free move area which can be seen in Figure 5.1 as the area outlined by white tape.

It was also observed in the video footage of participants while carrying out the experiment for the BG trials that during the control condition participants were more likely to move into the sweet spot while guessing the avatar's gaze direction towards PoI than they were during the TSB condition. In the video footage many participants can be seen moving away from the buckets during the control condition opposed to moving towards the buckets when the TSB condition was on. For an example of this see Figure 5.5 image *2-A* and *2-B* where the participant doing the control condition has moved directly from the starting position to the sweet spot. This indicates the reliance on the sweet spot when viewing a 2D display. However, judging by observations and the subsequent results obtained previously in *Study 1* and now in this empirical study, the benefits of the sweet spot only occur if the participant was standing in the sweet spot before and during the avatar's initial movement as it directed its gaze towards a PoI to initiate a trial. Ending up in the sweet spot subsequently to the avatar initiating the trial had no real advantage. This needs to be further investigated and verified in future work by doing a more extensive analysis of all the video footage captured from all fifteen participants in this empirical study (see Section 7.5).

The next chapter (see Chapter 6) outlines the third and final empirical study in this thesis. *Study 3* evaluates if the addition of 3D display technology to an avatar UI has any added advantage to the user's ability at interpreting the avatar's gaze direction.

Chapter 6

Empirical Study 3: '3D vs. TSB'

Considering the standard approach to displaying an avatar UI is to use a standard 2D display (i.e. which displays the avatar *monoscopically*) it is relatively easy and inexpensive to add 3D display technology using a software approach with the addition of cheap 3D glasses such as the anaglyph 3D glasses seen in Figure 6.1. It is reasonable to ask the question of whether or not the improvement in a participant's ability to interpret the avatar's gaze direction seen in the results of *Study 1* (see Section 4.4) and *Study 2* (see Section 5.3) during the TSB conditions can be matched, or surpassed, by using a standard 3D display technology (see Section 6.1). This empirical study answers this question directly and as such it is highlighted as a contribution of this thesis (see Section 1.2). This chapter is closely related to the '*Stereoscopic Avatar Interfaces: A study to determine what effect, if any, 3D technology has at increasing the interpretability of an avatar's gaze into the real-world*' (Dunne *et al.*, 2012a) publication mentioned in Section 1.4.

Figure 6.1: Anaglyph 3D glasses with filters in front of each eye: red filter covering the left eye and a cyan filter covering the right eye.

This empirical study presents the experimental set-up used to evaluate if 3D display technology has any bearing on a participant's ability to interpret the avatar's gaze direction, which is the first claim of the TSB technique (see Section 1.2). The task the participants had to carry out in this empirical study had the same layout *Study 1* (see Chapter 4), with seven floor markers in same positions. However, there is the addition of 3D display technology, increasing the two previous experimental conditions, *TSB* and *Control*, to four (i.e. *TSB*, *Control*, *3D TSB* and *3D Control*). Furthermore, the format of the experiment's procedure is more akin to *Study 2*. Unlike *Study 1* participants were not limited to standing on a particular floor marker while interpreting the avatar's gaze direction, they could freely move around the free move area and this is visible in Figure 5.2 as the area outlined by white tape. This means that each participants only had to do one trial per floor marker for each of the four experimental conditions, totalling twenty-eight trials split into two groups: fourteen trials with 3D and fourteen trials without 3D. Once again, the avatar behaved in exactly the same manner as it did in the two previous studies for all the experimental conditions. Also, participants started each of the trials either on the green or red square starting positions, these squares can be seen in Figure 6.2. For each group of fourteen trials the sequence was randomly chosen by the experimenter until all fourteen trials had been carried out, i.e. the avatar had directed its gaze towards each of the seven floor markers twice, once for the TSB condition and once for the Control condition. This process was obviously repeated for the second group of fourteen trials with the 3D technology turned on. Having the 3D technology turned on first or second was alternated between every second participant. The results of this experiment further support the findings seen in *Study 1* and *Study 2* in relation to the first claim of the TSB technique.

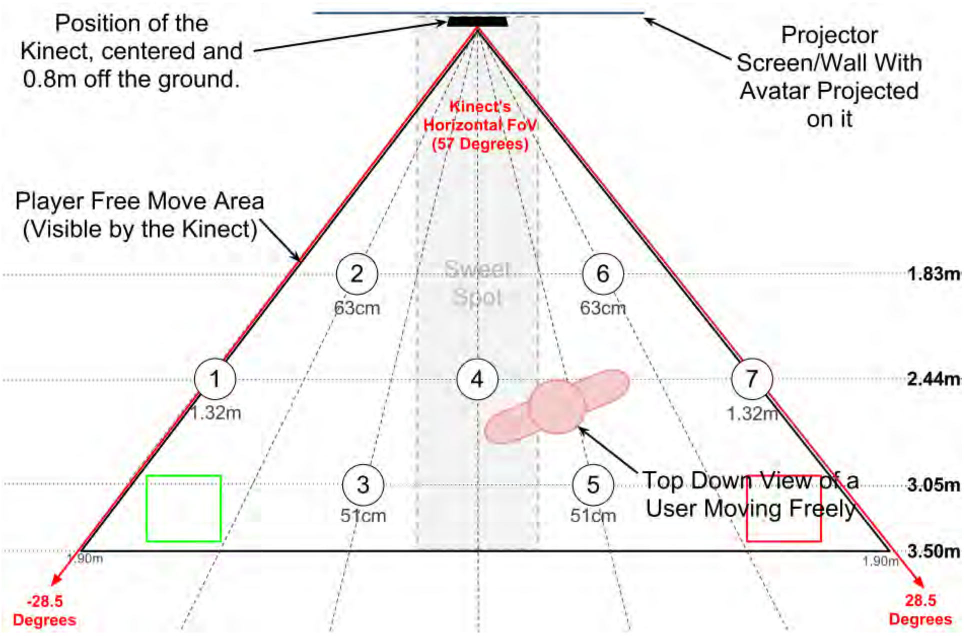


Figure 6.2: A floor plan for the '*3D vs. TSB*' experiment. Visible are the seven floor markers, the sweet spot, the two starting positions (i.e. green and red squares) and the FoV of the Kinect sensor (about 57°).

The next section will detail 3D technology and highlight the reason behind the choice of *Anaglyph 3D technology* for use in the experimental set-up in this empirical study (see Section 6.1). This section on 3D Technology is followed by Section 6.2 detailing the experimental procedure with a brief discussion on the participants used for this empirical study outlined in Section 6.3. The results are then presented in Section 6.4, which is followed by a discussion in Section 6.5 that concludes this empirical study, which is the final one of the series.

6.1 3D Technologies

There are many ways of achieving stereoscopic images on 2D displays, most require the user to wear filters (or lenses) over their right and left eyes. These filters can be *active* or *passive*, active filters are most commonly used for home entertainment system that use large LCD panel TV screens. These TV screens refresh the entire panel with an image rendered for the perspective of each eye, one at a time. The filter that corresponds to the image being displayed stays open, i.e. is transparent, while the other filter remains closed, i.e. is opaque. The images alternate on the screen between left and right eye images and the lenses also coincide with the images allowing the user to experience a fully stereoscopic image on the 2D display. Usually the glasses are synced to the TV through infra-red signals, ensuring that the correct lens is open at the right time.

In contrast to this, *passive* filter technology allows light into both eyes at the same time. However, the light that comes into each eye is only half the resolution of the 2D display. This is because the lenses are polarised, horizontally in one eye and vertically in the other, which then only allows light that is polarised by the same filter to pass through them. Compared to the *active* lenses, *passive* lenses are low cost, hence they are most commonly used in movie theatres.

Anaglyph 3D technology is also *passive* filter technology, however, anaglyph 3D glasses do not use polarised filters as it pre-dates the use of polarised *passive* filter technology, which are more common in movie theatres today. The filters in the anaglyph 3D glasses are instead coloured, one red and one blue (cyan), a different colour filter is placed over each of the viewer's eyes (i.e. red filter over the left eye, blue (cyan) filter over the right eye). The image is rendered so that slightly different perspectives of the image are produced for the left and right eyes, this is illustrated in Figure 6.3. The two coloured filters in the glasses match the colours in the rendered image. Hence, the rendered image for the right eye's perspective is coloured red so the red filter over the left eye blocks the red light from entering the left eye, the opposite occurs for the rendered image for the left eye's perspective. This means that the rendered image is interpreted by both of the viewer's eyes to be in the correct perspective for each eye at the same time and this is what causes the stereoscopic illusion that is created.

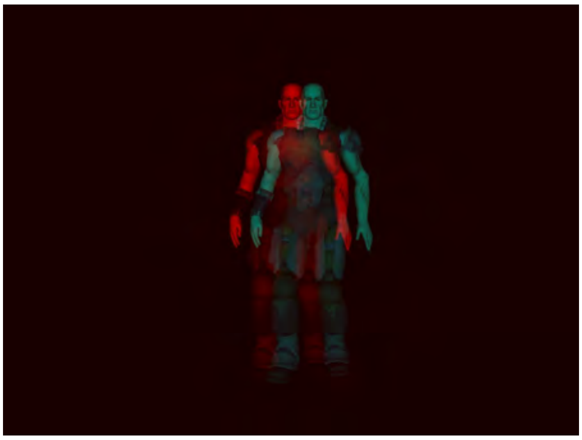


Figure 6.3: 3D model rendered with the anaglyph 3D stereoscopic effect.

Anaglyph 3D was chosen for this study as it can be easily used with a standard 2D display or projector with the addition of relatively low cost glasses that can be seen in Figure 6.1. It is easily switched on for 3D applications when a *NVIDIA*

GeForce GPU is installed along with the *NVIDIA 3D Vision Discover*[1] firmware. Anaglyph 3D is not a perfect solution as it can cause discomfort (He *et al.*, 2011) and cross-feedback or ghosting (Woods & Rourke, 2004) in some users. However, it still remains a complete and low cost solution to achieve stereoscopic 3D on a standard 2D display.

6.2 Procedure

The task designed for this experiment was similar to the task in *Study 1* with the '*42 Moves*' experiment. *Study 1* required the participants to interpret the avatar's gaze direction towards one of the floor markers (i.e. floor markers *1* to *7*) and make a guess which of the seven floor markers the avatar was looking at, forty-two times per experimental condition. In contrast to *Study 1*, this empirical study only required the participants to do one trial per floor marker for each of the four experimental conditions (*TSB*, *Control*, *3D TSB* and *3D Control*).

The participants did two groups of fourteen trials, one with the 3D display technology on and another group with the 3D technology off, both sets of fourteen were randomised. The random order was selected each time by the experimenter who selected from the list of fourteen (see Appendix C.1) possible trials per group at random until all fourteen of the trials had an outcome. Once again this was similar to the two previous studies, incorrect ('*0*') and correct ('*1*'). This randomised selection process was done twice, once for each of the two groups. Similar to *Study 2*, the two core experimental conditions (i.e. *TSB* and *Control*) that have remained consistent

[1] *NVIDIA 3D Vision Discover*. Anaglyph 3D, although not a perfect solution, is a complete and low cost way to achieve stereoscopic 3D. Website: `http://www.nvidia.co.uk/object/3D_Vision_Discover_Main_uk.html`

over all three empirical studies were mixed up during each set of fourteen trials, so the participant was constantly switching between the control condition and the TSB condition. The two sets of fourteen trials were conducted back to back and no survey was required as the quantitative data gathered was all that was required to evaluate if 3D display technology had any effect on the participants' abilities to interpret the avatar's gaze direction (i.e. the first claim of the TSB technique, see Section 1.2).

As this empirical study was carried out immediately after *Study 2*, participants were asked if they wanted to take part in this additional experiment with a duration of less than six minutes, if they agreed they waited in the holding area while the experimenter set up this new experiment. These participants did not have to sign another release form as the release form from *Study 2* was also suitable for this experiment. The experimenter placed the seven floor markers, labelled *1* to *7*, on the exact same positions of the seven PoI from the *Red* PoI group in *Study 2*, which were the exact same position as the original seven floor markers in *Study 1*. When the experimenter had completed the set-up, the participants would leave the holding area and stand in either the green or red square starting position. Every second participant started the first fourteen trials with the anaglyph 3D technology turned on, meaning they had to wear the anaglyph 3D glasses (see Figure 6.1). They were given an opportunity to try on the 3D glasses and get comfortable with the 3D effect on the 2D display. If participants were experiencing any ghosting (Woods & Rourke, 2004) the 3D effect was adjusted in the *NVIDIA 3D Vision Discover* firmware until the ghosting disappeared but they were still experiencing a stereoscopic image on the 2D display.

Each experimental trial began with the participant standing in one of two fixed starting positions, a green or red square, which are outlined in Figure 6.2. The participants would start on the opposite square after every two trials, e.g. red, red, green, green, red, red, green, green, etc., etc. A trial consisted of the avatar directing its gaze at the participant and then redirecting it to one of the seven floor markers once the participant had indicated that they were ready. Once the avatar had directed its gaze towards a floor marker, the participant was free to move within the *Free Move Area* while they interpreted the avatar's gaze direction. This is in contrast to *Study 1* where participant remained stationary while guessing and it replicated more natural style developed in *Study 2*, which allowed participants to move freely. The participant had to determine which of the seven floor markers the avatar was looking at and, when they were ready, would indicate their choice by speaking aloud the number of that floor marker. Once the participant's choice was recorded, the participant would return to either the green or red square starting position to begin the next trial, the experimenter told the participant which square they were to stand on.

Each of the seven participants performed twenty-eight trials. The avatar was set-up to look at each of the seven floor markers (see Figure 6.2 for the experiment layout) once for each of the four experimental conditions which were as follows:

- **Control**: The avatar appears as it would on a regular 2D display, i.e. the rendering does not update to reflect a participant's viewing position.
- **3D Control**: The same as **Control** but with anaglyph 3D switched on.

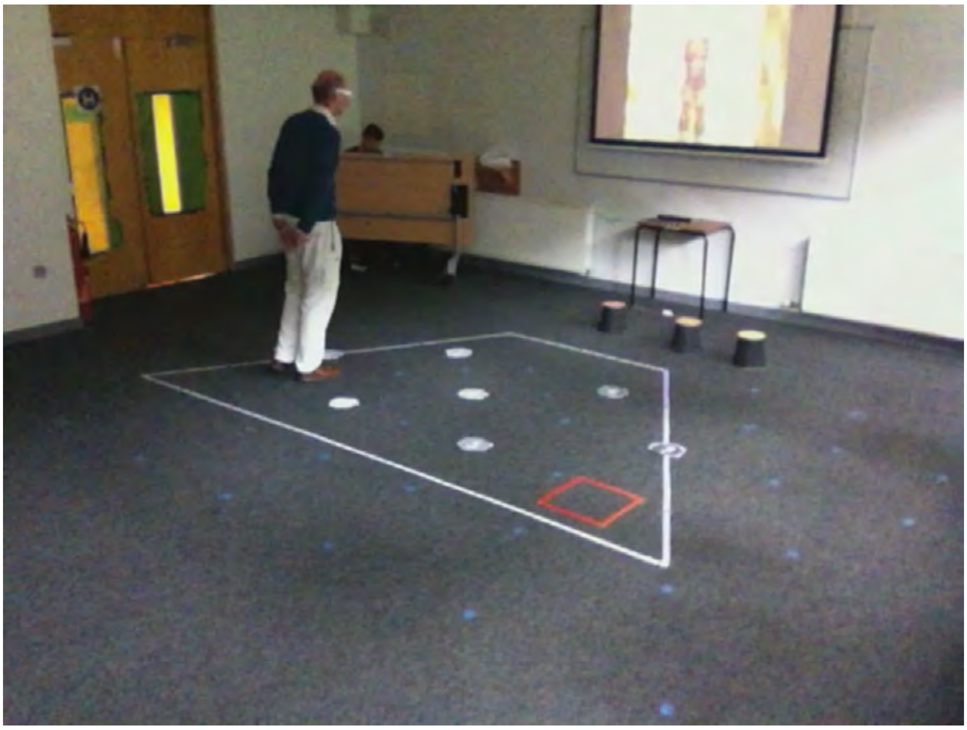


Figure 6.4: A participant performing the *3D vs. TSB* experiment wearing Anaglyph 3D glasses (the participant is standing behind floor marker *1*).

- **TSB**: The TSB technique is switched on, therefore the image of the 3D avatar is continuously updated to reflect the participant's perspective.

- **3D TSB**: The same as **TSB** above but with Anaglyph 3D switched on.

Learning effects were controlled by varying the sequence of the four experimental conditions across all participants. The sequence in which the floor markers were gazed at by the avatar were also varied between participants. This design resulted in each participant carrying out one trial per floor marker per condition (see Appendix C.1 for an example of the data collection sheet).

6.3 Participants

There were seven participants in total (two females, five males) with ages ranging from twenty-five to fifty-eight years. As each participant is different the settings for the 3D technology were adjusted to their individual comfort levels and tested before they began the experiment. The participant in Figure 6.4 is carrying out the experiment while wearing the anaglyph 3D glasses.

6.4 Results

Table 6.1 shows the mean accuracy rating achieved by all seven participants for each of the seven floor markers across all four experimental conditions. A *single factor ANOVA* statistical test was carried out to test for a significant difference between the four experimental conditions for each of the seven floor markers. The results at the $p < 0.05$ level [$F(3, 24) = 6.97$, $p = 0.0016$] show that a significant difference exists. This difference was further investigated through post hoc testing – in this case five *paired two sample for means Student t-Tests* were carried out on the results from Table 6.1 and are as follows:

1. *3D Control* (M=0.41, SD=0.28) and *TSB* (M=0.76, SD=0.07): $t(6)=-3.2$, $p < 0.01$. A statistically significant difference between the 3D and TSB results exists, with the latter out-performing the 3D condition. The test was done to compare 3D technology on a 2D display with the TSB technique on the same 2D display but without 3D technology turned on.

Table 6.1: The accuracy ratings for each of the seven floor markers as (see Figure 6.2) from all the data gathered from the seven participants is listed in this table below for all four experimental conditions (*TBS*, *TBS with 3D*, *control* and *control with 3D*).

Floor Marker	**Control**	**TSB**	**3D Control**	**3D TSB**
1	0.57	0.71	0.86	0.57
2	0.29	0.71	0.29	0.57
3	0.29	0.71	0.14	0.57
4	0.71	0.86	0.29	0.86
5	0.0	0.71	0.14	0.86
6	0.14	0.71	0.43	0.86
7	0.71	0.86	0.71	0.86
Mean	**0.39**	**0.76**	**0.41**	**0.73**

2. *3D Control* (M=0.41, SD=0.28) and *Control* (M=0.39, SD=0.28): t(6)=0.13, p=0.9. The control condition shows no statistically significant difference with 3D technology turned on and off.

3. *3D TSB* (M=0.73, SD=0.15) and *TSB* (M=0.76, SD=0.07): t(6)=-0.48, p=0.64. When the TSB technique is used there is no statistically significant difference with 3D technology turned on and off.

4. *3D TSB* (M=0.73, SD=0.15) and *3D Control* (M=0.41, SD=0.28): t(6)=-2.64, p=0.04. When the 3D technology is turned on with both the TSB technique and the control, a statistically significant difference exists between them (see Appendix C.2 for the complete set of results for each of the seven participants while 3D technology was on).

5. *TSB* (M=0.76, SD=0.07) and *Control* (M=0.39, SD=0.28): t(6)=3.39, $p < 0.01$. Likewise, when the 3D technology is turned off, there is a statistically significant difference between the TSB technique and the control.

The results of the first *paired two sample for means Student t-Test* (*3D Control* vs. *TSB*) indicate that the accuracy of gaze interpretations made by participants, when the TSB technique is used on a standard 2D display, outperforms the results when 3D display technology is used without the TSB technique. However, the results of the second *paired two sample for means Student t-Test* show no significant difference between the results for *3D Control* and *Control* conditions. This pattern is repeated in the third *paired two sample for means Student t-Test* when the comparison of the *3D TSB* and *TSB* conditions show no significant difference.

This outcome indicates that 3D technology alone does not eliminate the need for a technique like TSB when the avatar needs to be able to direct its attention to PoI in the viewer's real-world surroundings if the user is outside of the sweet spot. Furthermore, the addition of 3D display technology does not improve the effect achieved using the TSB technique, which clearly enables viewers to reliably interpret the avatar's gaze direction (i.e. indicated by the results of the fourth and fifth t-Tests). The next section will concluded this empirical study and end Chapter 6 with some final discussion on the findings.

6.5 Discussion

The results from the *3D vs. TSB* experiment when using 3D technology showed no significant effect on a participant's ability to accurately interpret an avatar's gaze. The results further back up the first claim of the TSB technique as detailed in Section 1.2. More specifically, when the avatar's gaze is being directed at a PoI in the user's real-world surroundings the TSB technique preforms just as well with or without 3D display technology (i.e. anaglyph 3D). For the specific experimental set-up in this empirical study there seems to have been no obvious benefit for the use of 3D display technology. Once again, the comparisons between both the TSB and control conditions show similar results to what has previously been seen in *Study 1* (see Section 4.4) and *Study 2* (see Section 5.3), where the TSB conditions returned higher accuracy ratings for participants as they interpreted the avatar's gaze direction.

The results of this empirical study clearly show that using 3D display technology (i.e. anaglyph 3D) in conjunction with an avatar UI, has little to no bearing on the user's ability to accurately interpret the avatar's gaze direction. Displaying an avatar UI to the user on a 2D display with 3D display technology enabled does not eliminate image distortion from the user's perspective caused by lateral foreshortening as they view the 2D display from outside of the sweet spot. At best, 3D display technology only enhances the way that the user perceives the image of the avatar – the avatar appears *stereoscopically* rather than *monoscopically*, which would have to make the avatar appear more realistic but only if viewed from the sweet spot or when the TSB technique is switched on. Further research (see Section 7.5) is needed to test if this is the case and a survey with specific questions targeting a participant level of perceived realism (see Section 2.1.1) would be needed. That being said, making a user wear an uncomfortable pair of 3D glasses in order to interact casually with an avatar UI on a 2D display could potentially negate any positive effects it brings in terms of realism in the first instance and this too will need to be addressed in future research.

Chapter 7

General Discussion

7.1 Overview

This thesis outlines the TSB technique, a graphical process that renders a 3D animated avatar onto a 2D display and continuously updates the rendered image to reflect the changing perspective of the user as they move during their interaction with the avatar. Based on the background literature (see Chapter 2), specifically research by Cuijpers *et al.* (2010); Eichner *et al.* (2007); Kipp & Gebhard (2008), three areas needed improvement: (1) not limiting the user to the sweet spot of the 2D display during interactions with an avatar UI, (2) increasing the user's perception of the avatar's presence and (3) tracking the user continuously in order to render the avatar correct for the user's perspective at all times. Our approach to rendering an avatar UI to a user on a 2D display was guided by these three areas with the development of the TSB technique with its combination of three graphical processes (i.e. *turning*, *stretching* and *boxing*). These three processes work in tandem with each other to deliver a consistent 3D illusion for the 3D avatar from the perspective of the user, without requiring the user to remain in the sweet spot. Head-tracking data capture from the Microsoft Kinect sensor is used to update the three graphical processes to create the sustained 3D illusion. The three graphical processes which make up the TSB technique are not novel per se. However, what is novel is their combined effect that helps to sustain the 3D illusion of the avatar that allows the user to accurately interpret the gaze direction of the avatar as it looks towards PoI in the user's real-world surroundings. This increase in interpretability of the avatar's gaze direction is due to the fact the avatar is corresponding geometrically correct (i.e. increased levels of Pl experienced by the user) to the user's real-world surroundings from the user's perspective and in turn this increases the user's perceived *corporeal* presence (Holz *et al.*, 2011) of the avatar. The avatar's representation also remains visually correct as the user moves freely in front of the 2D display, which increases the level of Psi experienced by the user.

Listed below are the two claims relating to the benefits of the using the TSB technique to render an avatar UI to a 2D display (see Section 1.2) and were both evaluated in this thesis through empirical studies (see Chapters 4 - first and second claims, Chapter 5: first claim and Chapter 6: first claim), the claims are as follows:

1. *Increase interpretability of the avatar's gaze direction*

2. *Increase perception of corporeal presence for the avatar*

The approach taken to evaluate the two claims of the TSB technique listed above are detailed in the next section (see Section 7.2), where a final discussion on each of the three empirical studies (see Chapters 4, 5 and 6) carried out in this thesis takes place consecutively. The three empirical studies form the major part of the contributions of this thesis (see Section 1.2). Then there is a discussion on some additional observations that occurred during the empirical studies in Section 7.3. This is followed by a discussion on a new approach needed to further evaluate the wall mounted PoI from the *Black* PoI group in *Study 2*. Followed by the topic of the Microsoft Kinect's range limitations although an update to Kinect SDK in November 2012 increased the range to greater than four metres. This increase in range means that more PoI could be placed in the participants' real-world surroundings and the free move area used in all three empirical studies could in fact be increased. This would have been a nice facility to have during the studies as it would have given participants more room to move within the free move area while interpreting the avatar gaze direction towards PoI. As well as that, the Kinect's FoV is about 57° and realistically a wider FoV would be needed to make it more usable in larger spaces. Ideally the FoV should be closer to 180°. Finally, Section 7.5 outlines planned future work with final analysis being shared in Section 7.6, bringing this thesis to an end.

7.2 Approach for Evaluating the TSB Technique

In order to evaluate the TSB technique fully, three empirical studies were devised (see Chapters 4, 5 and 6) and carried out as follows:

1. *Study 1* (see Chapter 4): This first empirical study evaluated both claims of the TSB technique (see Section 1.2). In *Study 1*, there was an increase in the participants' abilities to accurately interpret the avatar's gaze direction, which

was due to a higher level of Pl experienced by the participants. Consequently, an increase in Pl correlates with a higher level of Psi being experienced by the user and ultimately leads to an overall increase in corporeal presence. Furthermore, this could potentially increase the user's perceived presence of the avatar as a social entity as the user experiences higher levels Psi in the form of realism, immersion and social richness (Lombard & Ditton, 1997) during their interaction with the avatar. The experimental set-up used in *Study 1* tested the first and second claim of the TSB technique by having the participants do gaze perception trials, where they had to guess which floor marker the avatar was directing its gaze towards. There were seven floor markers and each participant was required to do forty-two trials guessing for each floor marker from every other floor marker, just once for each trial. Each participant had to carry out the forty-two trials twice once for the control condition and once for the TSB condition. The results of *Study 1* proceeded to prove both claims of the TSB technique. Overall the participants scored a higher accuracy rating at interpreting the avatar's gaze direction during the TSB condition and meant an increase Pl as the avatar was able to geometrically correspond more accurately to the PoI in the user's real-world surroundings from the user's own perspective. Subsequently, this increased the user's perceived corporeal presence of the avatar and correlated with the increase in Psi in the ratings of the survey that further indicated an increase in perceived corporeal presence for the avatar. Simply put, participants perceived a higher level of *corporeal* presence for the avatar during the TSB condition. Both the higher accuracy rating for participants' interpretability of the avatar's gaze direction during the trials and their subsequent higher ratings for the survey questions were statistically significant when the TSB technique was switched on.

2. *Study 2* (see Chapter 5): This second study pushed the boundaries of the scope in which the avatar could direct its gaze into the participants' real-world surroundings. The number of PoI placed throughout the laboratory increased by twenty-three to a total of thirty and they were thoroughly spread throughout the experimentation environment in four groups (*Red*, *Purple*, *Yellow* and *Black*). Overall, the broader scope of the PoI in *Study 2* meant that the TSB condition and the control condition were pushed to their limit in order to see if the results from *Study 1* could be replicated. *Study 2* unlike *Study 1* forced participants to start each trial in either the green or red square starting positions, this was decided after the findings in *Study 1* highlighted that participants preformed equally well for both experimental conditions when trials started from the sweet spot. Once the avatar had directed its gaze towards a PoI the participants were allowed to move from their starting square to anywhere within the free move area in order to guess what PoI the avatar was directing its gaze towards. This ensured participants started each trial from a non-optimal viewing angle. Observations of participants in the recorded video footage showed that they naturally moved towards the sweet spot to help them guess during the control condition.

   However, the results indicate that participants did not get any benefit from being in the sweet spot after the avatar had already directed its gaze. Furthermore, it would seem that any benefit from being in the sweet spot would only occur if they were standing in the sweet spot prior to the avatar directing its gaze towards a PoI. The results of the second study backed up the findings of *Study 1* in relation to the first claim of the TSB technique, as there was a higher accuracy rating achieved by participants during the TSB condition and these higher accuracy ratings were significantly different from the accuracy ratings achieved by participants during the control condition. It would seem from these results that the TSB technique is useful and works in large spaces by allowing the avatar rendered to the 2D display to indicate towards specific PoI in a user's immediate vicinity. Also, it is important to note that it may be useful to carry out another investigation into wall mounted PoI which are seen in the *Black* PoI group. The results indicated that during the control condition participants were as likely to guess correctly when interpreting the avatar's gaze direction as they were during the TSB condition. There was no immediately obvious explanation for these results during the initial analysis of the quantitative data gathered. It is possible that the LG trials were too easy and they may have benefited from being surrounded by more letters placed on the walls as distraction PoI, as the placement of the actual eight PoI may have been too systematic, making the results predictable. This group of PoI are important, as the interpretation of this gaze direction behind a participant at eye level usually indicates the presence of someone else approaching from behind or is a natural gaze behaviour to engage in when giving directions to someone, i.e. to occasionally look in the direction of the destination you are directing someone to during an interaction.

3. *Study 3* (see Chapter 6): To conclude, the third and final empirical study was carried out to evaluate if the first claim of the TSB technique with the addition 3D display technology (i.e. anaglyph 3D) could match, or surpass, the findings from *Study 1*. Where participants generally achieved higher accuracy ratings at interpreting the avatar's gaze direction during the TSB condition on a regular 2D display (i.e. without 3D display technology). In *Study 3* the participants were required to carry out twenty-eight trials, two groups of fourteen trials one with 3D being used and the other without 3D. Within each of these groups of fourteen trials the participant did seven trials for the TSB condition and seven for the control condition. The result of this empirical study mirrored those of *Study 1* and *Study 2* where there were statistically significant differences between the experimental conditions, with the TSB condition out preforming the control condition. There was no significant benefit to a participant's ability to accurately interpret the avatar's gaze direction when 3D display technology was switched on.

It is evident that any additional benefit derived from using 3D display technology with an avatar UI is purely aesthetic, taking the rendered image of the avatar from *monoscopic* to *stereoscopic* and adding a higher level of realism to the rendered image. *Realism* is an important conceptualisation of presence according to Lombard & Ditton (1997). Yet, users are required to wear 3D glasses which can be impractical for real-world scenarios as well as being uncomfortable to wear for extended periods of time. In contrast to this, the current glasses-free 3D technology (i.e. *autostereoscopic*) suffers from restrictive viewing angles but it is realistic to assume that an avatar UI would be suited to the *autostereoscopic* technology when it advances out of its infancy. Autostereoscopic technology's major advantage would be a realistic 3D rendering without the need for uncomfortable and somewhat impractical 3D glasses, leading to a more immersive user experience. This needs further research and is discussed again in more detail in the future work section (see Section 7.5).

Considering the insignificant differences between the results of the TSB conditions (*TSB* and *3D TSB*) from *Study 3*, a user based study with a clear emphasis on evaluating the level of perceived *corporeal* presence for an avatar for both TSB conditions could be a fruitful endeavour and a useful piece of *future work* (see Section 7.5).

7.3 Additional Findings and Observations

In addition to the evaluation of the TSB technique detailed in Chapter 1 Section 1.2, the exploration and quantifying of the limits of 2D display technology were conducted. This was done in order to ensure a fair and unbiased comparison with a standard 2D display technology and a 2D display with the TSB technique switched on. This was achieved by ensuring that throughout each of the three empirical studies (see Chapters 4, 5 and 6) the 2D display used during the control condition trials for each participant was the same 2D display used during the TSB condition trials with the only exception being that the latter had the TSB technique switched on. This ensured that each of the participants experienced both conditions with the same 2D display. Furthermore, the use of the same 2D display for both conditions for all participants ensured unbiased and fair results.

An investigation of the sweet spot in *Study 1* shows there is little or no effect on the user's ability to accurately interpret the avatar's gaze direction when the TSB condition is switched on or off (see Section 4.4). However, the results from *Study 2* (see Chapter 5 Section 5.3) show this is only true if the user was standing in the sweet spot before the avatar directed its gaze towards a PoI. Moving to the sweet spot after the avatar has directed its gaze towards the PoI does not increase the user ability to accurately interpret the gaze direction when the TSB technique is not used.

In *Study 2* (see Chapter 5) the further a participant moved away from the 2D display and more specifically the Kinect, the lower the accuracy achieved at interpreting the avatar's gaze direction. This decrease was consistent across all participants during the experimental conditions that used the Kinect, i.e. when the TSB technique was switched on. The cause could be the angle of the Kinect, which was placed directly under the 2D display and was pointing slightly upwards. This meant, as the user moved further away from the Kinect they were perceived to get shorter in height. This had a knock-on effect to their perspective of the avatar as the avatar looked out towards PoI in the user's real-world surroundings.

7.4 Limitations of the *TSB Technique*

One limitation to the TSB technique is how people or other user's with their own avatars will perceive another user's avatar on the 2D display as it is rendered to match that other user's perspective. As it stands the TSB technique provides an interface that is one-to-one, i.e. one avatar per user. However, there may be a potential problem when many users have avatars within the same environment and an investigation into what the impact of these additional renderings of avatars will have on users within that environment who may or may not have their own avatar will be important. Especially as users' viewing angles line up or are exactly opposite to each other. Having a personalised avatar for each user will obviously help users determine their own avatar as they move throughout an environment with other users' avatars. Quite possibly in busy environments with multiple users and their avatars, some avatars may appear perspectively correct for a than the intended user at any one time. For non-tracked users it may become distracting if they occasionally feel as if they are being engaged with by an avatar only to find out the avatar was looking behind them. Previous research (Agrawala *et al.*, 1997; Kulik *et al.*, 2011) used polarised lens to deliver separate images to multiple users, however, it is not practical to make people wear 3D glasses on the off chance of needing to interact with an avatar. To counter act this users would most likely receive audio through a headset attached to their mobile device and any dialogue the user would have would be picked up on a microphone attached to their mobile device also.

Another limitation of the TSB technique's current set-up is the visibility of the avatar from a user's perspective when they are viewing the avatar from a very acute viewing angle. From an acute viewing angle the box of the boxing process occludes the avatar and in order for the user to see the avatar from these angles the avatar needs to step forward towards the front of the box, i.e. the notional window of the box. As it stands the occlusion of the avatar by the boxing process encourages users to step towards the sweet spot of the 2D display. This limitation off course impacts the avatar's ability to communicate using gestures such as pointing as well as other movements when the avatar is occluded by the box. The current set-up of the TSB technique accounts for the FoV of the Kinect along with the width and height of the 2D display used in all the three empirical studies. A greater FoV for the Kinect would mean that the user could be tracked at these more acute viewing angles, helping to counteract lateral foreshortening over a wider area. A larger 2D display would simply mean a the more space to render the avatar to and the box in which the avatar is placed could be bigger (i.e. height and width), meaning the avatar would remain visible at more acute viewing angles before occlusion from the boxing process would occur. This along with a wider FoV on the Kinect would allow a tracked user to view the avatar in the correct perspective for them across a much larger area in front of the 2D display. These two factors,the Kinect's FoV and the 2D display size, can counteract this current limitation.

7.5 Future Research

Any future research with regard to further evaluation or development of the TSB technique would culminate in several areas:

- **Further investigation into presence**: The increase seen so far in Pl, Psi and the overall *corporeal* presence perceived by participants indicates that further research is needed. The next experiment should focus more on the subjective experiences of the participants in relation to presence and a more extensive survey will be carried out to determine a participant's level of perceived presence for the avatar. Using the lessons learned in all three empirical studies (see Chapters 4, 5 and 6) from this thesis, where participants clearly gained higher accuracy ratings when the TSB technique was used, a far more elaborate experimental approach to test for increases in participants' perceived presence due to the increases in gaze perception, should be devised and carried out. It will be important to establish a baseline for real corporeal presence, this would be achieved by the addition of two extra experimental conditions. These two additional experimental conditions would precede the TSB and control conditions seen in the three empirical studies in this thesis.

The first extra experimental condition would be to use a real person (i.e. an actor) to play the role of the avatar as seen in all three of the previous empirical studies, by having the actor stand in front of the participant with their back to the 2D display on which the avatar will eventually appear in the subsequent experimental conditions. In the case of *Study 1* (see Chapter 4) an actor would proceed by following a script (i.e. one of either path **1** through **7**) and look towards the floor markers in order to facilitate the participants' guesses. The actor would have to carry out the role exactly as the avatar would, avoiding talking and pointing. The second of the extra experimental conditions would be to record the actor doing all seven path scripts on video and play these recorded videos to participants accordingly. However, this approach would require a lot of preparation in advance of carrying out the experiment. Then the participant would progress onto doing the two original experimental conditions with the avatar on the 2D display like the previous empirical studies. This would establish a real sense of perceived *corporeal* presence for an actual human actor in the minds of the participants and would help ground the subjective nature of corporeal presence when participants are being asked to rate their experience in a survey. In addition, the use of *biometric* measurements could be useful to evaluate a participant's physiological responses to perceived corporeal presence but a baseline would have to be measured first, which again would require a human actor.

- **Further presence work with 3D technology**: Considering the insignificant difference between the results of the TSB conditions (*TSB* and *3D TSB*) in *Study 3*, with and without 3D technology being switched on, a user based study with a clear emphasis on evaluating the level of perceived presence for an avatar for both TSB conditions is a good next step. Keeping in mind the many difficulties in evaluating presence due to its subjective nature and the inadequacies of using surveys to gather qualitative data outlined by Holz *et al.* (2011) in this regard, a robust experimental design would be required.

- **Multiple users**: An investigation into the use of the TSB technique with multiple users and how it will be achieved; if the users will be effected by each other's avatars when they appear on 2D displays from the other users' perspectives (i.e. turned, stretched and boxed) is needed. This is an obvious and important next step if the research in this thesis is ever going to be used in actual real-world scenarios.

- **Avatar transitioning between displays**: Another interesting area for future research lies in the area of how to transition avatars between 2D displays in a user's real-world surroundings. If the avatar is going to appear on one 2D display why not have it appear on multiple 2D displays or better yet, have the avatar follow or guide the user by transitioning from 2D display to 2D display that happen to be scattered throughout the user's real-world surroundings.

- **A 180° Kinect rig**: With a recent update to the Kinect SDK (November 2012) it is now possible to run multiple Kinects on a single machine, meaning that placing a least five Kinects in a semi-circle in front of the display could provide up to 180° of FoV. This in addition to the increased range of the Kinect to in excess of four metres, and means the laboratory space used to conducted all three empirical studies in this thesis could have been utilised entirely, thus allowing the user's to freely move anywhere within the space.

- **Further analysis of '*Study 2*' video footage**: Initial early analysis of the video footage of participants recorded during *Study 2* experiments (see Chapter 5) highlighted a recurring phenomenon, where participants who moved to the sweet spot after the avatar had already directed its gaze towards a PoI garnered no advantage from being in the sweet spot. It would seem the benefits of the sweet spot only occur if the participant was standing in the sweet spot before and during the avatar's initial movement as it directed its gaze towards a PoI. This needs further investigation and a planned evaluation of all recorded footage of participants carrying out *Study 2* experiments where the phenomenon was first noticed is required.

- **Does 3D technology increase realism**: In *Study 3* the addition of 3D display technology had no effect on a participants' abilities to accurately interpret the avatar's gaze direction. However, it is possible that where 3D technology fails to increase the participants' abilities to interpret the avatar's gaze, it may compensate for this by increasing the participant's perceived Psi and subsequently, realism (see Section 2.1.1). This would then have to be considered as a means to increase the perceived corporeal presence of an avatar UI in its real-world surroundings. A carefully designed survey will be needed to evaluate this claim, as it is a possibility that the requirement to wear 3D glasses with most 3D technology may negatively effect the Psi in the first place. This would need careful consideration and if possible, the use of glasses-free 3D technology (i.e. autostereoscopy technology) if available, may remove any issues of wearing 3D glasses.

- **Incorporation of deictic gestures**: The incorporation of deictic gestures, such as pointing, could be an interesting development direction to take with the TSB technique. Pointing where you are directing your gaze is a natural reaction for a human interlocutor when referencing an object to another person. It would be interesting to carry out experimentation to test if the addition of accurate deictic gestures could effect the accuracy rating of a participant's ability to interpret the avatar's gaze direction.

7.6 Final Thoughts

The three empirical studies detailed in this thesis evaluated the TSB technique's use in conjunction with an avatar UI on a 2D display, and the findings of these empirical studies form the basis of three of the major contributions of this thesis in addition to the TSB technique itself and they are as follows:

1. The use of the TSB technique with an avatar UI rendered on a 2D display, where the user is free to move outside of the sweet spot without losing the 3D illusion of the rendered avatar and subsequently, increasing the perceived level of corporeal presence for the avatar in the user's real-world surroundings.

2. Evaluation to test the effect of the TSB technique on a user's ability to interpret the avatar's gaze direction (see Chapters 4, 5 and 6).

3. Evaluation to test the effect of the TSB technique on a user's perception of the avatar's corporeal presence (see Chapter 4).

4. Evaluation to test if 3D display technology has any bearing on a user's ability to interpret the avatar's gaze direction with and without the TSB technique being switched on (see Chapter 6).

The quantitative results from the above three empirical studies are pretty definitive and support the first claim of the TSB technique, indicating that the TSB technique generally increases a user's accuracy rating while interpreting the avatar's gaze direction towards PoI in the user's real-world surroundings. This can be attributed to the fact that the 3D illusion of the avatar on the 2D display is maintained by the TSB technique from the participant's perspective at all times, as they are tracked by the Kinect. This allows for the avatar to geometrically correspond correctly to its real-world surroundings from the user's perspective. This increase has also been shown to have a positive affect on the participant's sense of the avatar '*being there*' with them, which Slater (2009) refers to as Pl where the avatar is geometrically corresponding correctly to its real-world surroundings from the participant's perspective. This in turn represents an increase in the participant's experienced level of Psi and this ultimately correlates in an observable increase in perceived *corporeal* presence (Holz *et al.*, 2011) for the avatar and backs up the second claim of the TSB technique.

Furthermore, the qualitative survey results used in *Study 1* indicates that participants did experience increases in Psi during the TSB condition. This increase in Psi also indicates that the participants were able to suspend their disbelief more readily, which is an important trait in mediated communication. Where the human communicator does not feel as if they are talking to the medium (i.e. projection on a 2D display), as this can prevent a natural communication style from occurring.

Also, the use of 3D display technology has no bearing on a user's ability to interpret the avatar's gaze direction with or without the TSB conditions being switched on. The results of *Study 3* (see Chapter 6) detail another contribution of this thesis (see Section 1.2) and show that there was no additional benefit to participants while they interpreted the avatar's gaze direction (i.e. the first claim of the TSB technique) when 3D display technology (i.e. anaglyph 3D) was used. In fact, it is further evidence that a graphical process such as the TSB technique is still required with all 2D displays, regardless of 3D technology being used, to give users the freedom to move and increase their ability to interpret the avatar's gaze direction from outside the sweet spot for the 2D display.

The observed increase in Pl in all three empirical studies and Psi in the survey carried out in *Study 1*, contributes to an increase in perceived *corporeal* presence when the TSB technique is switched on, which can only be seen as a benefit for avatar interfaces. It could also be argued from the results of quantitative analysis in *Study 1* that the TSB technique definitely increases Pl and the resulting effect is that the avatar geometrically corresponds correctly to its real-world surroundings from a user's perspective. In turn, this could be the reason behind the observable increase of Psi as perceived by the participants, which is what the results of the survey in *Study 1* indicated with increased ratings for realism, immersion and social richness.

Other ways to increase *corporeal* presence could be the use of more realistic 3D models and animations, advanced dialogue systems, adding human-like behaviours (i.e. gestures and gaze), etc. (Andrist *et al.*, 2012; Jan *et al.*, 2009; Lim *et al.*, 2009; Mao & Gratch, 2009; McQuiggan & Lester, 2007; Miksatko *et al.*, 2010; Ochs & Sabouret, 2009; Rickel & Johnson, 1999; Rushforth *et al.*, 2009; Steptoe & Steed, 2008; Swartout *et al.*, 2006). However, as the TSB technique is a relatively easy graphical process to implement within a 3D graphical engine with the relative ease of integrating the Microsoft Kinect sensor, it could be an easy fit for other research, such as Andrist *et al.* (2012); Steptoe & Steed (2008), who both have developed high-fidelity gaze behaviour models for avatar interfaces. Consequently, they may find it useful to augment the realism achieved by their high-fidelity gaze behaviour models with the TSB technique in order to increase the perceived *corporeal* presence of their avatars. Also, as the TSB technique allows for better geometric correspondence between the avatar and the user's real-world surroundings from the user's perspective, the user can interpret the avatar's gaze direction from outside of the sweet spot. This is where the addition of the TSB technique or a similar graphical process could be most useful and *a direction worth looking towards.*

Appendix A

Chapter 4 Additional Material

Research Participant Release Form

You have been asked to be a participant in Mark Dunne's ongoing research within the LOK8 (locate) research team in Kevin Street, Dublin Institute of Technology (DIT). You may be recorded on video during the experiment, however, this recording and any answers (Name, Age, Gender, Job Description) you will be asked to give will only be used for the purposes of this research. Which means that your data and the data collected from other participants will contribute to the overall findings for this ongoing research, and will be used in future academic publications.

The LOK8 research team will endeavour to do the following:

- To respect the individuals freedom to decline participation.
- To maintain confidentiality of research data.
- To be responsible for maintaining ethical standards.
- To NOT specifically identify individuals with their data (i.e. use of a video or photo) unless it is necessary, and then only after the individual has given consent.
- To take every precaution and make every effort to minimise potential risk to participants safety and that of their data.
- To only use the data supplied by the participant with their full consent.

"I hereby give my consent to the Mark Dunne and LOK8 research team in DIT to use the data gathered from my participation in this experiment and any video recordings of myself for purposes of the ongoing research and future publications only."

Name (block capitals): ______________________________

Date: _____/_____/2011

Signature: ______________________________

Office Use Only

Participant No: __________

1st Condition: _____ ***& 1st Journey:*** _____

2nd Condition: _____ ***& 2nd Journey:*** _____

Figure A.1: Each of the thirty-one participants were required to sign this release form before being instructed to carry out the experiment.

Participant No.	Sex	Age	First Condition	TSB %	Control %	Difference %
1	m	27	TSB	57.14%	47.62%	9.52%
2	m	26	Control	71.43%	54.76%	16.67%
3	m	29	TSB	71.43%	30.95%	40.48%
4	m	28	Control	66.67%	45.24%	21.43%
5	m	64	TSB	71.43%	30.95%	40.48%
6	m	33	Control	69.05%	47.62%	21.43%
7	m	26	TSB	59.52%	54.76%	4.76%
8	m	33	Control	50.00%	57.14%	-7.14%
9	m	23	TSB	85.71%	30.95%	54.76%
10	m	30	Control	76.19%	40.48%	35.71%
11	f	24	TSB	66.67%	40.48%	26.19%
12	m	25	Control	61.90%	38.10%	23.81%
13	f	26	TSB	73.81%	33.33%	40.48%
14	m	57	Control	85.71%	50.00%	35.71%
15	f	57	TSB	69.05%	33.33%	35.71%
16	m	33	Control	59.52%	33.33%	26.19%
17	m	29	TSB	78.57%	42.86%	35.71%
18	f	25	Control	61.90%	33.33%	28.57%
19	f	27	TSB	57.14%	28.57%	28.57%
20	m	39	Control	78.57%	47.62%	30.95%
21	m	39	TSB	45.24%	47.62%	-2.38%
22	f	10	Control	64.29%	47.62%	16.67%
23	m	6	TSB	54.76%	35.71%	19.05%
24	f	35	Control	76.19%	45.24%	30.95%
25	f	61	TSB	71.43%	33.33%	38.10%
26	m	23	Control	73.81%	61.90%	11.90%
27	m	24	TSB	40.48%	26.19%	14.29%
28	m	32	Control	66.67%	38.10%	28.57%
29	f	27	TSB	66.67%	40.48%	26.19%
30	m	25	Control	64.29%	47.62%	16.67%
31	f	26	TSB	52.38%	38.10%	14.29%
			Average %	**66.05%**	**41.40%**	**24.65%**

Figure A.2: Complete list of results for each of the thirty-one participant in *Study 1*: Mean accuracy rating (%) for the *TSB Condition*, mean accuracy rating (%) for the *Control Condition* and the difference between both conditions.

All Possible Moves (42)	Random Set 1 (42)	Random Set 2 (42)	Random Set 3 (42)	Random Set 4 (42)	Random Set 5 (42)	Random Set 6 (42)	Random Set 7 (42)
1 - 2	1 - 6	2 - 6	3 - 7	4 - 7	5 - 2	6 - 3	7 - 6
1 - 3	6 - 7	6 - 2	7 - 4	7 - 3	2 - 5	3 - 7	6 - 7
1 - 4	7 - 6	2 - 7	4 - 7	3 - 4	5 - 4	7 - 4	7 - 5
1 - 5	6 - 2	7 - 2	7 - 5	4 - 2	4 - 7	4 - 5	5 - 1
1 - 6	2 - 4	2 - 4	5 - 6	2 - 6	7 - 4	5 - 6	1 - 6
1 - 7	4 - 5	4 - 7	6 - 7	6 - 2	4 - 6	6 - 2	6 - 5
2 - 1	5 - 2	7 - 3	7 - 1	2 - 1	6 - 5	2 - 5	5 - 3
2 - 3	2 - 1	3 - 6	1 - 4	1 - 6	5 - 3	5 - 3	3 - 7
2 - 4	1 - 5	6 - 4	4 - 2	6 - 7	3 - 6	3 - 4	7 - 1
2 - 5	5 - 6	4 - 2	2 - 5	7 - 2	6 - 2	4 - 6	1 - 5
2 - 6	6 - 4	2 - 3	5 - 1	2 - 3	2 - 1	6 - 5	5 - 7
2 - 7	4 - 3	3 - 5	1 - 5	3 - 7	1 - 2	5 - 4	7 - 2
3 - 1	3 - 2	5 - 2	5 - 7	7 - 4	2 - 3	4 - 2	2 - 4
3 - 2	2 - 7	2 - 5	7 - 2	4 - 6	3 - 1	2 - 1	4 - 3
3 - 4	7 - 3	5 - 6	2 - 6	6 - 5	1 - 6	1 - 3	3 - 5
3 - 5	3 - 4	6 - 1	6 - 5	5 - 1	6 - 3	3 - 2	5 - 6
3 - 6	4 - 2	1 - 5	5 - 2	1 - 4	3 - 7	2 - 4	6 - 4
3 - 7	2 - 5	5 - 1	2 - 7	4 - 5	7 - 3	4 - 3	4 - 6
4 - 1	5 - 1	1 - 4	7 - 6	5 - 6	3 - 2	3 - 5	6 - 1
4 - 2	1 - 4	4 - 3	6 - 4	6 - 1	2 - 7	5 - 2	1 - 4
4 - 3	4 - 1	3 - 7	4 - 6	1 - 7	7 - 1	2 - 3	4 - 7
4 - 5	1 - 7	7 - 6	6 - 3	7 - 1	1 - 5	3 - 6	7 - 4
4 - 6	7 - 4	6 - 7	3 - 5	1 - 5	5 - 6	6 - 4	4 - 1
4 - 7	4 - 6	7 - 1	5 - 3	5 - 4	6 - 4	4 - 7	1 - 2
5 - 1	6 - 5	1 - 3	3 - 6	4 - 1	4 - 2	7 - 5	2 - 6
5 - 2	5 - 4	3 - 4	6 - 1	1 - 3	2 - 6	5 - 7	6 - 3
5 - 3	4 - 7	4 - 5	1 - 2	3 - 6	6 - 7	7 - 3	3 - 2
5 - 4	7 - 1	5 - 4	2 - 1	6 - 3	7 - 6	3 - 1	2 - 1
5 - 6	1 - 2	4 - 1	1 - 6	3 - 5	6 - 1	1 - 5	1 - 3
5 - 7	2 - 6	1 - 6	6 - 2	5 - 3	1 - 3	5 - 1	3 - 6
6 - 1	6 - 1	6 - 5	2 - 4	3 - 2	3 - 4	1 - 7	6 - 2
6 - 2	1 - 3	5 - 7	4 - 3	2 - 4	4 - 1	7 - 2	2 - 3
6 - 3	3 - 6	7 - 4	3 - 4	4 - 3	1 - 4	2 - 6	3 - 1
6 - 4	6 - 3	4 - 6	4 - 5	3 - 1	4 - 5	6 - 1	1 - 7
6 - 5	3 - 5	6 - 3	5 - 4	1 - 2	5 - 1	1 - 6	7 - 3
6 - 7	5 - 3	3 - 1	4 - 1	2 - 5	1 - 7	6 - 7	3 - 4
7 - 1	3 - 7	1 - 2	1 - 7	5 - 7	7 - 5	7 - 1	4 - 2
7 - 2	7 - 5	2 - 1	7 - 3	7 - 5	5 - 7	1 - 4	2 - 5
7 - 3	5 - 7	1 - 7	3 - 1	5 - 2	7 - 2	4 - 1	5 - 4
7 - 4	7 - 2	7 - 5	1 - 3	2 - 7	2 - 4	1 - 2	4 - 5
7 - 5	2 - 3	5 - 3	3 - 2	7 - 6	4 - 3	2 - 7	5 - 2
7 - 6	3 - 1	3 - 2	2 - 3	6 - 4	3 - 5	7 - 6	2 - 7

Figure A.3: All seven of the random paths assigned to participants, each participant was randomly assigned two of the seven, one for each of the experimental conditions.

Non-Sweet Spot Symeterical Differences							
Accuracy (%) of Correct Moves							
Moves	**L-Control**	**Moves**	**R-Control**	**Moves**	**L-TSB**	**Moves**	**R-TSB**
2 - 1	100.00%	6 - 7	100.00%	2 - 1	89.66%	6 - 7	79.31%
3 - 1	75.86%	5 - 7	72.41%	3 - 1	89.66%	5 - 7	55.17%
5 - 1	13.79%	3 - 7	10.34%	5 - 1	58.62%	3 - 7	58.62%
6 - 1	24.14%	2 - 7	10.34%	6 - 1	72.41%	2 - 7	65.52%
7 - 1	10.34%	1 - 7	3.45%	7 - 1	75.86%	1 - 7	65.52%
1 - 2	89.66%	7 - 6	93.10%	1 - 2	96.55%	7 - 6	86.21%
3 - 2	34.48%	5 - 6	48.28%	3 - 2	86.21%	5 - 6	86.21%
5 - 2	55.17%	3 - 6	48.28%	5 - 2	82.76%	3 - 6	75.86%
6 - 2	55.17%	2 - 6	44.83%	6 - 2	75.86%	2 - 6	82.76%
7 - 2	24.14%	1 - 6	24.14%	7 - 2	55.17%	1 - 6	65.52%
1 - 3	10.34%	7 - 5	20.69%	1 - 3	34.48%	7 - 5	31.03%
2 - 3	20.69%	6 - 5	6.90%	2 - 3	31.03%	6 - 5	51.72%
5 - 3	58.62%	3 - 5	58.62%	5 - 3	48.28%	3 - 5	51.72%
6 - 3	10.34%	2 - 5	20.69%	6 - 3	48.28%	2 - 5	41.38%
7 - 3	0.00%	1 - 5	6.90%	7 - 3	65.52%	1 - 5	65.52%
N	15		15		15		15
AVERAGE	38.85%		37.93%		67.36%		64.14%
Standard Deviation	0.32		0.32		0.21		0.16
Student T-Test		**0.69**				**0.33**	
Correlation		0.96				0.80	
Sweet Spot Only Symeterical Differences							
Accuracy (%) of Correct Moves							
Moves	**L-Control**	**Moves**	**R-Control**	**Moves**	**L-TSB**	**Moves**	**R-TSB**
4 - 1	58.62%	4 - 7	44.83%	4 - 1	41.38%	4 - 7	27.59%
4 - 2	48.28%	4 - 6	62.07%	4 - 2	89.66%	4 - 6	89.66%
4 - 3	65.52%	4 - 5	48.28%	4 - 3	62.07%	4 - 5	44.83%
N	3		3		3		3
AVERAGE	57.47%		51.72%		64.37%		54.02%
Standard Deviation	0.09		0.09		0.24		0.32
Student T-Test		**0.62**				**0.19**	
Correlation		-0.83				0.99	

Figure A.4: This table contains the results of two Student t-Tests between the right and the left sides of the floor markers for both conditions, one test excluding the sweet spot and another with the sweet spot only. The results of the tests show there is no statistical difference between either of the side for both conditions for each *paired two sample for means Student t-Test.*

Participant No.	Sex	Age	First Condition	To what degree did you become so involved in doing the task that you lost all track of time?	To what degree did you feel the 3D virtual character's head movements were natural?	To what degree did you feel the 3D virtual character's gaze direction towards the spots on the ground was realistic?	To what degree did you feel the 3D virtual character was responsive to your actions?	To what degree did your experience with the 3D virtual character's gaze seem consistent with your real world experiences?	To what degree do you think the 3D virtual character was actually able to look out at the real world 'spots' on the ground?
1	m	27	TSB	3	4	2	5	4	2
2	m	26	Control	4	3	3	4	4	3
3	m	29	TSB	4	3	2	4	2	2
4	m	28	Control	2	3	3	4	2	3
5	m	64	TSB	4	3	2	1	3	3
6	m	33	Control	4	3	2	5	2	2
7	m	26	TSB	2	2	2	2	1	2
8	m	33	Control	5	4	2	3	1	2
9	m	23	TSB	5	3	4	5	5	3
10	m	30	Control	3	4	4	4	3	3
11	f	24	TSB	4	3	2	3	2	2
12	m	25	Control	2	3	1	2	1	1
13	f	26	TSB	3	2	2	3	2	1
14	m	57	Control	3	3	3	2	2	3
15	f	57	TSB	4	4	3	2	2	3
16	m	33	Control	4	5	4	5	3	3
17	m	29	TSB	2	2	1	2	1	1
18	f	25	Control	4	3	4	3	2	3
19	f	27	TSB	5	2	2	3	2	2
20	m	39	Control	3	4	4	1	1	3
21	m	39	TSB	4	4	5	5	4	4
22	f	10	Control	4	3	1	4	1	1
23	m	6	TSB	3	1	1	4	1	1
24	f	35	Control	5	3	3	4	3	5
25	f	61	TSB	3	3	2	1	3	2
26	m	23	Control	2	4	3	5	4	4
27	m	24	TSB	5	3	4	4	3	5
28	m	32	Control	3	4	3	1	3	4
29	f	27	TSB	4	5	4	4	5	4
30	m	25	Control	4	5	3	3	3	4
31	f	26	TSB	4	3	3	5	4	3
			Average Score	**3.58**	**3.26**	**2.71**	**3.32**	**2.55**	**2.71**

Figure A.5: List of all the survey results for the *control condition* from each of the thirty-one participants in *Study 1*.

Participant No.	Sex	Age	First Condition	To what degree did you become so involved in doing the task that you lost all track of time?	To what degree did you feel the 3D virtual character's head movements were natural?	To what degree did you feel the 3D virtual character's gaze direction towards the spots on the ground was realistic?	To what degree did you feel the 3D virtual character was responsive to your actions?	To what degree did your experience with the 3D virtual character's gaze seem consistent with your real world experiences?	To what degree do you think the 3D virtual character was actually able to look out at the real world 'spots' on the ground?
1	m	27	TSB	4	5	5	5	4	2
2	m	26	Control	5	5	4	4	4	4
3	m	29	TSB	4	3	3	4	4	3
4	m	28	Control	3	3	4	3	4	4
5	m	64	TSB	5	4	3	1	4	4
6	m	33	Control	4	4	4	3	2	3
7	m	26	TSB	4	3	4	5	4	3
8	m	33	Control	3	4	3	2	3	2
9	m	23	TSB	3	5	4	5	3	5
10	m	30	Control	3	5	4	4	4	4
11	f	24	TSB	2	3	3	5	3	3
12	m	25	Control	3	4	4	2	3	4
13	f	26	TSB	4	5	3	5	4	3
14	m	57	Control	1	3	4	2	2	4
15	f	57	TSB	5	4	3	5	4	4
16	m	33	Control	4	5	5	4	4	4
17	m	29	TSB	3	4	4	3	2	1
18	f	25	Control	5	4	4	4	3	4
19	f	27	TSB	5	3	3	3	3	3
20	m	39	Control	4	5	5	4	4	4
21	m	39	TSB	4	4	5	5	5	4
22	f	10	Control	5	4	5	3	4	5
23	m	6	TSB	4	3	3	4	3	3
24	f	35	Control	5	3	3	4	3	3
25	f	61	TSB	3	4	4	5	4	4
26	m	23	Control	2	2	4	4	4	4
27	m	24	TSB	5	4	2	4	4	5
28	m	32	Control	4	4	4	4	4	4
29	f	27	TSB	3	4	4	5	4	4
30	m	25	Control	3	4	3	4	3	3
31	f	26	TSB	4	5	3	4	4	4
			Average Score	**3.74**	**3.94**	**3.74**	**3.84**	**3.55**	**3.58**

Figure A.6: List of all the survey results for the *TSB condition* from each of the thirty-one participants in *Study 1*.

Appendix B

Chapter 5 Additional Material

	Path Number	1			
	Participant Number				
No.	Move Number	Condition	Control Score (0/1)	TSB Score (0/1)	Placement
1	4(8)	B			W1S
2	21(4)	A			N2W
3	16(1)	A			N2E
4	10(1)	B			N1W
5	F	B			
6	A	A			
7	19(7)	B			N3E
8	1(1)	A			E2S
9	3(5)	A			N1E
10	14(3)	B			E2S
11	B	B			
12	18(5)	A			W3S
13	17(3)	B			E1S
14	11(4)	A			W2S
15	2(4)	B			N1W
16	G	A			
17	5(2)	B			W1S
18	8(7)	A			N2W
19	6(2)	A			N3E
20	20(2)	B			N3W
21	C	A			
22	D	B			
23	12(7)	A			N2E
24	13(6)	B			E3S
25	11(4)	B			N1E
26	15(5)	A			E2S
27	H	B			
28	9(3)	A			W3S
29	22(6)	A			E3S
30	21(4)	B			W3S
31	7(6)	B			N3W
32	E	A			
	Totals		/16	/16	

Figure B.1: *Path 1* data sheet, shows a randomised order for the bucket guessing (BG) trials (see Chapter 5) which are done in sets of four trials and broken up with a wall letter guessing trials. Each of the three buckets are placed in the same positions for each trial for all the users to ensure consistency in the results. This is achieved with a compass styled placement scheme as seen in Figure 5.4.

	Path Number	2			
	Participant Number				
No.	Move Number	Condition	Control Score (0/1)	TSB Score (0/1)	Placement
1	4(8)	A			W1S
2	21(4)	B			N2W
3	16(1)	B			N2E
4	10(1)	A			N1W
5	F	A			
6	A	B			
7	19(7)	A			N3E
8	1(1)	B			E2S
9	3(5)	B			N1E
10	14(3)	A			E2S
11	B	A			
12	18(5)	B			W3S
13	17(3)	A			E1S
14	11(4)	B			W2S
15	2(4)	A			N1W
16	G	B			
17	5(2)	A			W1S
18	8(7)	B			N2W
19	6(2)	B			N3E
20	20(2)	A			N3W
21	C	B			
22	D	A			
23	12(7)	B			N2E
24	13(6)	A			E3S
25	11(4)	A			N1E
26	15(5)	B			E2S
27	H	A			
28	9(3)	B			W3S
29	22(6)	B			E3S
30	21(4)	A			W3S
31	7(6)	A			N3W
32	E	B			
	Totals		/16	/16	

Figure B.2: *Path 2* data sheet is identical to the *Path 1* data sheet but the conditions are reversed for each trial.

Participant Number: _____

"Found It" Experiment

Research Participant Release Form

You have been asked to be a participant in Mark Dunne's ongoing research within the LOK8 (locate) team in Kevin Street, Dublin Institute of Technology (DIT). You may be recorded on video during the experiment, however, video recording, recorded data and any information (Name, Age, Gender, Job Description) you will be asked to give will only be used for the purposes of this research. That means the data you provide will contribute to the overall findings for the LOK8 research, and will be used in academic publications.

The LOK8 research team will endeavour to do the following:

- To respect the individual's freedom to decline participation.
- To maintain confidentiality of research data.
- To be responsible for maintaining ethical standards.
- To NOT specifically identify individuals with their data (i.e. use of a video or photo) unless it is necessary, and then only after the individual has given consent.
- To take every precaution and make every effort to minimise potential risk to participants safety and that of their data.
- To only use the the data supplied by the participant with their full consent.

"I hereby give my consent to the Mark Dunne and the LOK8 research team in DIT to use the data gathered from my participation in this experiment, including any video recordings of me for purposes of their ongoing research and future publications ***only****:"*

Name (block capitals): ______________________________

Sex (M/F): ______________________________

Age: ______________________________

Job Title/Description: ______________________________

Contact Email Address: ______________________________

Signature: ______________________________

Date: _____/_____/2012

Figure B.3: Each of the fifteen participants were required to sign this release form before being instructed to carry out the study.

Participant Number	Path	Condition	1(1)	5(2)	9(3)	2(4)	3(5)	13(6)	8(7)	4(8)	10(1)	6(2)	14(3)	11(4)	15(5)	7(6)	12(7)	16(1)	20(2)	17(3)	21(4)	18(5)	22(6)	19(7)	A	B	C	D	E	F	G	H	*Average*
1	1	Control	1		0		1		1			0		0	0		0	0			0	1	0		1		0		0		0		
3	1	Control	1		0		0		0			1		1	0		0	0			0	1	0		1		1		1		1		
5	1	Control	1		0		1		0			0		1	0		0	0			0	0	0		0		0		0		1		
7	1	Control	0		0		1		1			0		1	0		0	0			0	0	0		0		1		0		0		
9	1	Control	0		0		1		0			0		1	1		0	0			0	0	0		1		1		0		1		
11	1	Control	0		0		1		0			1		0	0		1	0			0	0	0		0		0		1		0		
13	1	Control	1		0		0		0			0		1	0		0	0			0	0	0		0		1		1		1		
15	1	Control	0		0		0		0			0		0	0		0	0			0	0	0		1		0		1		1		
			50%		0%		63%		25%			25%		63%	13%		13%	0%			0%	25%	0%		50%		50%		50%		63%		30.47%
2	2	Control		1		1		0		1	0		1	1		0			0	0	0			1		1		1		1		1	
4	2	Control		0		0		1		0	0		0	0		0			1	0	1			1		0		0		0		1	
6	2	Control		1		1		1		0	1		0	0		1			0	0	0			0		1		1		0		0	
8	2	Control		1		1		0		1	1		0	1		1			1	1	0			1		0		1		1		0	
10	2	Control		1		1		0		0	1		1	1		1			0	0	0			0		1		1		0		1	
12	2	Control		1		0		1		1	1		0	1		1			1	0	1			1		0		0		0		1	
14	2	Control		1		0		1		1	0		1	0		0			0	0	0			1		0		0		1		1	
				86%		57%		57%		57%	57%		43%	57%		57%			43%	14%	29%			71%		43%		57%		43%		71%	52.68%
1	1	TSB		1		1		1		1	1		0	1		1			1	1	1			1		1		1		1		0	
3	1	TSB		1		1		1		1	1		0	0		1			0	1	1			1		1		1		0		1	
5	1	TSB		1		1		1		1	1		1	1		1			1	1	0			0		1		1		1		0	
7	1	TSB		1		1		1		1	1		0	1		1			1	1	1			0		1		1		1		1	
9	1	TSB		1		1		1		1	0		1	1		1			1	1	0			0		1		1		1		0	
11	1	TSB		1		1		1		1	1		1	1		1			0	1	1			0		1		1		1		0	
13	1	TSB		1		1		1		1	0		1	1		1			1	1	1			1		1		1		1		1	
15	1	TSB		1		1		0		1	1		1	1		1			1	1	0			0		1		1		1		0	
				100%		100%		88%		100%	75%		63%	88%		100%			75%	100%	63%			38%		100%		100%		88%		38%	82.03%
2	2	TSB	1		1		1		1			1		1	1		1	1			0	0	0		1		1		1		1		
4	2	TSB	1		1		1		1			0		1	0		1	0			0	1	1		0		0		1		0		
6	2	TSB	0		1		0		0			0		0	0		0	1			0	1	1		0		1		0		0		
8	2	TSB	0		1		1		1			1		0	1		1	1			0	0	1		1		1		1		1		
10	2	TSB	1		1		1		0			1		1	1		1	0			0	0	1		0		1		1		1		
12	2	TSB	1		1		1		1			1		1	1		1	0			0	1	0		1		1		1		1		
14	2	TSB	1		1		0		0			0		1	1		1	1			1	1	1		1		1		0		1		
			71%		100%		71%		57%			57%		71%	71%		86%	57%			14%	57%	71%		57%		86%		71%		71%		66.96%

Figure B.4: This is the complete set of results for each of the fifteen participants doing trial for all thirty PoI in *Study 2* (see Section 5.3 of Chapter 5) for both the *TSB* condition and the *control condition*.

Appendix C

Chapter 6 Additional Material

	Participant Number			
No.	**Move Number**	**Condition**	**Control Score (0/1)**	**TSB Score (0/1)**
4	11(4)	A		
12	15(5)	B		
1	10(1)	A		
10	14(3)	B		
2	6(2)	A		
8	10(1)	B		
6	7(6)	A		
14	12(7)	B		
3	14(3)	A		
11	11(4)	B		
5	15(5)	A		
9	6(2)	B		
7	12(7)	A		
13	7(6)	B		
	Totals		/7	/7

Figure C.1: A sample of a data collection sheet for the *Study 3*: '*3D vs. TSB*' (see Chapter 6). The '*No.*' column represents the fourteen moves which have been randomly reordered. The '*Move Number*' column contains the set of seven moves times two, ensuring both conditions are done for each move. '**A**' is the *control condition* and '**B**' is the *TSB condition.*

Sequence Number		1	2	3	4	5	6	7	8	9	10	11	12	13	14	
Participant Number	Condition	10(1)	6(2)	14(3)	11(4)	15(5)	7(6)	12(7)	10(1)	6(2)	14(3)	11(4)	15(5)	7(6)	12(7)	Totals
9	A	1	0	0	1	0	1	1								4
9	B								1	1	1	1	1	1	1	7
10	A	1	1	0	1	0	1	0								4
10	B								1	1	1	1	1	1	1	7
11	A	1	0	0	0	0	0	1								2
11	B								1	1	1	1	1	1	1	7
12	A	1	0	1	0	1	1	1								5
12	B								0	0	0	1	1	1	1	4
13	A	1	0	0	0	0	0	1								2
13	B								0	0	0	1	1	1	1	4
14	A	0	0	0	0	0	0	0								0
14	B								0	1	0	0	1	0	0	2
15	A	1	1	0	0	0	0	1								3
15	B								1	0	1	1	0	1	1	5
Totals		6	2	1	2	1	3	5	4	4	4	6	6	6	6	

Figure C.2: This is the complete set of results for each of the seven participants while wearing the 3D (anaglyph) glasses in *Study 3* (see Chapter 6). '**A**' is the *3D control condition* and '**B**' is the *3D TSB condition*.

References

Agrawala, M., Beers, A.C., McDowall, I., Fröhlich, B., Bolas, M. & Hanrahan, P. (1997). The two-user Responsive Workbench: support for collaboration through individual views of a shared space. In *International Conference on Computer Graphics and Interactive Techniques: Proceedings of the 24 th annual conference on Computer graphics and interactive techniques*, 327–332, pp. xiii, 11, 12, 145

Al Moubayed, S. & Skantze, G. (2011). Effects of 2d and 3d displays on turn-taking behavior in multiparty human-computer dialog. In *Proceedings of the 15th Workshop on the Semantics and Pragmatics of Dialogue (SemDial'11)*, vol. 192–193, pp. 51

Al Moubayed, S., Alexanderson, S., Beskow, J. & Granström, B. (2011). A robotic head using projected animated faces. In *Proceedings of the International Conference on Audio-Visual Speech Processing 2011*, 71, pp. 51

Al Moubayed, S., Edlund, J. & Beskow, J. (2012). Taming mona lisa: communicating gaze faithfully in 2d and 3d facial projections. *ACM Transactions on Interactive Intelligent Systems (TiiS)*, **1**, 11, pp. 43

Andrist, S., Pejsa, T., Mutlu, B. & Gleicher, M. (2012). A head-eye coordination model for animating gaze shifts of virtual characters. In *Proceedings of the 4th Workshop on Eye Gaze in Intelligent Human Machine Interaction*, 4, ACM, pp. 153

Argyle, M. & Cook, M. (1976). Gaze and mutual gaze. *Cambridge U Press*, **XI**, 210, pp. 44

Argyle, M. & Dean, J. (1965). Eye-contact, distance and affiliation. *Sociometry*, **28**, 289–304, pp. 35, 44

Argyle, M., Ingham, R., Alkema, F. & McCallin, M. (1973). The different functions of gaze. *Semiotica*, **7**, 19–32, pp. 43

Babu, S., Suma, E., Barnes, T. & Hodges, L.F. (2007). Can immersive virtual humans teach social conversational protocols? In *Virtual Reality Conference, 2007. VR'07. IEEE*, 215–218, IEEE, pp. 33, 38, 40

Bailenson, J., Beall, A., Blascovich, J., Raimundo, M. & Weisbuch, M. (2001). Intelligent agents who wear your face: Users' reactions to the virtual self. In *Intelligent Virtual Agents*, 86–99, Springer, pp. 35, 36, 38, 40

Bates, J. *et al.* (1994). The Role of Emotion in Believable Agents. *Communications of the ACM*, **37**, 122–125, pp. 29

Baylor, A. & Kim, Y. (2005). Simulating instructional roles through pedagogical agents. *International Journal of Artificial Intelligence in Education*, **15**, 95–115, pp. 29

Baylor, A.L. (2005). The impact of pedagogical agent image on affective outcomes. In *International Conference on Intelligent User Interfaces, San Diego, CA*, pp. 29

Beskow, J. & Al Moubayed, S. (2010). Perception of gaze direction in 2d and 3d facial projections. In *Proceedings of the ACM/SSPNET 2nd International Symposium on Facial Analysis and Animation*, 24–24, ACM, pp. 43

Biocca, F., Harms, C., Burgoon, J.K., Interface, M. & Lansing, E. (2003). Towards A More Robust Theory and Measure of Social Presence: Review and Suggested Criteria. *Presence*, **12**, 456–480, pp. 24, 38

Bosse, T., Gerritsen, C. & Treur, J. (2007). Integration of rational and biological factors in virtual agent decision making. In *ICCM'07*, Citeseer, pp. 27

Carpenter, M., Tomasello, M., Wetherby, A. & Prizant, B. (2000). Joint attention, cultural learning, and language acquisition: Implications for children with autism. *Autism spectrum disorders: A transactional developmental perspective*, 31–54, pp. 44, 50

Cassell, J., Stocky, T., Bickmore, T., Gao, Y. & Nakano, Y. (2002). Mack: Media lab autonomous conversational kiosk. In *Proc. of Imagina*, vol. 2, 12–15, pp. xi, 1, 3, 4, 7, 27, 38, 40

Choi, Y.K., Miracle, G.E. & Biocca, F. (2001). The effects of anthropomorphic agents on advertising effectiveness and the mediating role of presence. *Journal of Interactive Advertising*, **2**, pp. 34

Cruz-Neira, C., Sandin, D. & DeFanti, T. (1993). Surround-screen projection-based virtual reality: the design and implementation of the CAVE. In *Proceedings of the 20th annual conference on Computer graphics and interactive techniques (CGIT'93)*, 135–142, ACM, pp. 32, 42

Cuijpers, R., van der Pol, D. & Meesters, L. (2010). Mediated eye-contact is determined by relative pupil position within the sclera. In *Proceedings of the European Conference on Visual Perception (ECVP'10)*, vol. 129, pp. 46, 48, 49, 50, 52, 53, 136

Dahlbäck, N., Jönsson, A. & Ahrenberg, L. (1993). Wizard of oz studies—why and how. *Knowledge-based systems*, **6**, 258–266, pp. xx, 75, 101, 107

De Freitas, S., Rebolledo-Mendez, G., Liarokapis, F., Magoulas, G. & Poulovassilis, A. (2010). Learning as immersive experiences: Using the four-dimensional framework for designing and evaluating immersive learning experiences in a virtual world. *British Journal of Educational Technology*, **41**, 69–85, pp. 41

Dohi, H., Ishizuka, M. & Engineering, C. (2009). A life-like agent interface system with second life avatars on the opensimulator server. In *Intelligent Virtual Agents*, 182–190, Springer Verlag, Amsterdam, The Netherlands, pp. 42

DUNNE, M., MAC NAMEE, B. & KELLEHER, J. (2009). Intelligent virtual agent: Creating a multi-modal 3d avatar interface. In *Proceedings of the 9th Annual Information Technology & Telecommunication Conference (IT&T 2009)*, 24, pp. 20

DUNNE, M., MAC NAMEE, B. & KELLEHER, J. (2010a). Scalable multi-modal avatar interface for multi-user environments. In *The International Conference on Computer Animation and Social Agents 2010 (CASA 2010)*, Saint Malo, France, pp. 20

DUNNE, M., MAC NAMEE, B. & KELLEHER, J. (2010b). Tsb technique: Increasing a user's sense of immersion with intelligent virtual agents. In *The 21st National Conference on Artificial Intelligence and Cognitive Science Student Symposium (AICS 2010)*, Galway, Ireland, pp. 20, 21

DUNNE, M., MAC NAMEE, B. & KELLEHER, J. (2012a). Stereoscopic avatar interfaces: A study to determine what effect, if any, 3d technology has at increasing the interpretability of an avatar's gaze into the real world. In *Multimodal Analyses enabling Artificial Agents in Human-Machine Interaction (MA3-2012)*, Santa Cruz, Ca, USA, pp. 16, 20, 21, 123

DUNNE, M., MAC NAMEE, B. & KELLEHER, J. (2012b). The turning, stretching and boxing technique: A step in the right direction. In Y. Nakano, M. Neff, A. Paiva & M. Walker, eds., *Intelligent Virtual Agents*, vol. 7502 of *Lecture Notes in Computer Science*, 363–369, Springer Berlin Heidelberg, pp. 14, 15, 20, 21, 70

EICHNER, T., PRENDINGER, H., ANDRÉ, E. & ISHIZUKA, M. (2007). Attentive presentation agents. In *Intelligent Virtual Agents*, 283–295, Springer, pp. 46, 53, 136

GARAU, M., SLATER, M., BEE, S. & SASSE, M.A. (2001). The impact of eye gaze on communication using humanoid avatars. In *Proceedings of the SIGCHI conference on Human factors in computing systems*, 309–316, ACM, pp. 2

GEBHARD, P., SCHRÖDER, M., CHARFUELAN, M., ENDRES, C., KIPP, M., PAMMI, S., RUMPLER, M. & TÜRK, O. (2008). Ideas4games: building expressive virtual characters for computer games. In *Intelligent Virtual Agents*, 426–440, Springer, pp. 27, 42

GIANNACHI, G., GILLIES, M., KAYE, N. & SWAPP, D. (2009). Mediating performance through virtual agents. In *Intelligent Virtual Agents*, 439–445, Springer, pp. 32, 38

GIBSON, J. & PICK, A. (1963). Perception of another person's looking behavior. *The American journal of psychology*, **76**, 386–394, pp. 105

GOLDMAN, A.I. (2008). *Simulating Minds: The philosophy, psychology, and neuroscience of mindreading*. Oxford University Press, pp. 29

GRATCH, J., OKHMATOVSKAIA, A., LAMOTHE, F., MARSELLA, S., MORALES, M., VAN DER WERF, R. & MORENCY, L. (2006). Virtual rapport. In *Intelligent Virtual Agents*, 14–27, Springer, pp. 29

Gratch, J., Wang, N., Gerten, J., Fast, E. & Duffy, R. (2007). Creating rapport with virtual agents. In *Intelligent Virtual Agents*, 125–138, Springer, pp. 29

Hancock, M., Nacenta, M., Gutwin, C. & Carpendale, S. (2009). The effects of changing projection geometry on the interpretation of 3d orientation on tabletops. In *Proceedings of the ACM International Conference on Interactive Tabletops and Surfaces*, 157–164, ACM, pp. xiii, 12

He, S., Zhang, T. & Doyen, D. (2011). Visual discomfort prediction for stereo contents. In *Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series*, vol. 7863, 64–73, SPIE '11, pp. 128

Heerink, M., Kröse, B., Evers, V. & Wielinga, B. (2010). Relating conversational expressiveness to social presence and acceptance of an assistive social robot. *Virtual reality*, **14**, 77–84, pp. 24

Holz, T., Dragone, M. & O'Hare, G.M.P. (2008). Where Robots and Virtual Agents Meet. *International Journal of Social Robotics*, **1**, 83–93, pp. 27, 28

Holz, T., Campbell, A., O'Hare, G., Stafford, J., Martin, A. & Dragone, M. (2011). Mira—mixed reality agents. *International journal of human-computer studies*, **69**, 251–268, pp. iv, xii, 2, 4, 5, 6, 7, 17, 23, 26, 39, 68, 71, 72, 74, 80, 90, 94, 98, 137, 148, 152

Jan, D., Roque, A., Leuski, A., Morie, J. & Traum, D. (2009). A virtual tour guide for virtual worlds. In *Intelligent Virtual Agents*, 372–378, Springer, pp. 27, 153

Kipp, M. & Gebhard, P. (2008). Igaze: Studying reactive gaze behavior in semi-immersive human-avatar interactions. In *Intelligent Virtual Agents*, 191–199, Springer, pp. 11, 29, 30, 33, 38, 40, 42, 46, 50, 51, 53, 66, 136

Kita, S. (2003). Pointing: A foundational building block of human communication. *Pointing: Where Language, Culture, and Cognition Meet, Lawrence Erlbaum Associates, Inc., Mahwah, New Jersey*, **1**, 1–8, pp. 2

Kleinke, C. (1986). Gaze and eye contact: a research review. *Psychological bulletin*, **100**, 78–100, pp. 44

Koenderink, J., van Doorn, A., Kappers, A. & Todd, J. (2004). Pointing out of the picture. *Perception, London*, **33**, 513–530, pp. 47

Kruppa, M., Spassova, L. & Schmitz, M. (2005). The Virtual Room Inhabitant - Intuitive Interaction with Intelligent Environments. In S. Zhang & R. Jarvis, eds., *Australian Conference on Artificial Intelligence*, vol. 3809, 225–234, pp. 7, 38, 39

Kulik, A., Kunert, A., Beck, S., Reichel, R., Blach, R., Zink, A. & Froehlich, B. (2011). C1x6: a stereoscopic six-user display for co-located collaboration in shared virtual environments. *ACM Trans. Graph.*, **30**, 188:1–188:12, pp. 11, 12, 145

Langton, S. (2000). The mutual influence of gaze and head orientation in the analysis of social attention direction. *The Quarterly Journal of Experimental Psychology: Section A*, **53**, 825–845, pp. 43

Lim, M., Aylett, R., Ho, W., Enz, S. & Vargas, P. (2009). A socially-aware memory for companion agents. In *Intelligent Virtual Agents*, 20–26, Springer, pp. 153

Lobmaier, J.S., Fischer, M.H. & Schwaninger, A. (2006). Objects capture perceived gaze direction. *Experimental Psychology*, **53**, 117, pp. 2

Lombard, M. & Ditton, T. (1997). At the heart of it all: The concept of presence. *Journal of Computer-Mediated Communication*, **3**, 0–0, pp. xii, 2, 4, 5, 6, 8, 10, 15, 17, 22, 23, 25, 27, 30, 31, 34, 35, 36, 37, 38, 39, 41, 42, 67, 72, 74, 80, 84, 89, 98, 139, 142

Lombard, M., Ditton, T.B., Crane, D., Davis, B., Gil-Egui, G., Horvath, K. & Rossman, J. (2000). Measuring presence: A literature-based approach to the development of a standardized paper-and-pencil instrument. In *Third International Workshop on Presence, Delft, The Netherlands*, vol. 240, Citeseer, pp. 2, 4, 14, 25, 39

Mac Namee, B. & Dunne, M. (2009). Widening the evaluation net. In Z. Ruttkay, M. Kipp, A. Nijholt & H. Vilhjálmsson, eds., *Intelligent Virtual Agents*, vol. 5773 of *Lecture Notes in Computer Science*, 525–526, Springer Berlin Heidelberg, pp. 20

Mac Namee, B.M. & Kelleher, J.D. (2009). Stepping Off the Stage. In *Proceeding of the 22nd Annual Conference on Computer Animation and Social Agents (CASA '09)*, pp. xiv, 38, 39, 41

Mao, W. & Gratch, J. (2009). Modeling social inference in virtual agents. *AI & society*, **24**, 5–11, pp. 153

McCauley, L. & D'Mello, S. (2006). Miki: A speech enabled intelligent kiosk. In *Intelligent Virtual Agents*, 132–144, Springer, pp. xi, 1, 3, 4, 7, 38, 40

McQuiggan, S. & Lester, J. (2007). Modeling and evaluating empathy in embodied companion agents. *International Journal of Human-Computer Studies*, **65**, 348–360, pp. 153

McQuiggan, S., Rowe, J. & Lester, J. (2008). The effects of empathetic virtual characters on presence in narrative-centered learning environments. In *Proceedings of the twenty-sixth annual SIGCHI conference on Human factors in computing systems*, 1511–1520, ACM, pp. 31

Mehrabian, A. (1981). Silent messages: Implicit communication of emotion and attitudes. *Belmont, CA: Wadsworth*, pp. 36

Miksatko, J., Kipp, K. & Kipp, M. (2010). The persona zero-effect: evaluating virtual character benefits on a learning task with repeated interactions. In *Intelligent Virtual Agents*, 475–481, Springer, pp. 153

Mirenda, P., Donnellan, A. & Yoder, D. (1983). Gaze behavior: A new look at an old problem. *Journal of Autism and Developmental Disorders*, **13**, 397–409, pp. 43

Monk, A. & Gale, C. (2002). A look is worth a thousand words: full gaze awareness in video-mediated conversation. *Discourse Processes*, **33**, 257–278, pp. 14, 44

MORI, M. (1970). The uncanny valley. *Energy*, **7**, 33–35, pp. 27

MORI, M., MACDORMAN, K.F. & KAGEKI, N. (2012). The uncanny valley [from the field]. *IEEE Robot. Automat. Mag.*, **19**, 98–100, pp. 27

MUMME, C., PINKWART, N. & LOLL, F. (2009). Design and implementation of a virtual salesclerk. In *Intelligent Virtual Agents*, 379–385, Springer, pp. 27

NAKAMURA, J. & CSIKSZENTMIHALYI, M. (2009). *Flow Theory and Research*, chap. 18, 195–206. OUP USA, pp. 32

NORMAN, D.A. (2007). *The Design of Future Things*. Nueva York, EUA : Basic Books, pp. 24

OBDRZALEK, S., KURILLO, G., OFLI, F., BAJCSY, R., SETO, E., JIMISON, H. & PAVEL, M. (2012). Accuracy and robustness of kinect pose estimation in the context of coaching of elderly population. In *Engineering in Medicine and Biology Society (EMBC), 2012 Annual International Conference of the IEEE*, 1188–1193, IEEE, pp. 11

OCHS, M. & SABOURET, N. (2009). Simulation of the dynamics of virtual characters' emotions and social relations. In *Affective Computing and Intelligent Interaction and Workshops, 2009. ACII 2009. 3rd International Conference on*, 1–6, IEEE, pp. 153

O'HARE, G., DUFFY, B. & CAMPBELL, A. (2004). Nexus: Mixed reality experiments with embodied intentional agents. In *Proceedings of Computer Animation and Social Agents-CASA 2004*, pp. 39

Pan, X., Gillies, M. & Slater, M. (2008). Male bodily responses during an interaction with a virtual woman. In *Intelligent virtual agents*, 89–96, Springer, pp. 32, 36, 38

Perner, J. (1999). Theory of mind. *Developmental psychology: Achievements and prospects*, 205–230, pp. 25

Peters, C., Pelachaud, C., Bevacqua, E., Mancini, M. & Poggi, I. (2005). A model of attention and interest using gaze behavior. In *Intelligent Virtual Agents*, 229–240, Springer, pp. 9, 44, 45, 50, 51, 52

Peters, C., Castellano, G. & de Freitas, S. (2009). An exploration of user engagement in HCI. In *Proceedings of the International Workshop on Affective-Aware Virtual Agents and Social Robots - AFFINE '09*, 1–3, ACM Press, New York, New York, USA, pp. 32

Poggi, I., Pelachaud, C. & De Rosis, F. (2000). Eye communication in a conversational 3D synthetic agent. *AI communications*, **13**, 169–182, pp. 43, 45, 51, 52

Pol, v.d.D.D., Cuijpers, R., Meesters, L. *et al.* (2009). *The effect of slant on the perception of eye-contact*. Master's thesis, Technische Universiteit Eindhoven, pp. 47, 48, 49

Raskar, R., Brown, M.S., Yang, R., Chen, W.C., Welch, G., Towles, H., Scales, B. & Fuchs, H. (1999). Multi-Projector Displays Using Camera-Based Registration. In *IEEE Visualization*, 161–168, IEEE, pp. iv, 8

RationalCraft.com (2010). Winscape: Virtual Windows. *Retrieved June 2010.* Website: `http://www.rationalcraft.com/Winscape.html`, pp. 43, 57

Reeves, B. & Nass, C. (1996). *The media equation: How people treat computers, television, and new media like real people and places.*. Chicago, IL, US: Center for the Study of Language and Information; New York, NY, US: Cambridge University Press, pp. 24

Rettie, R. (2003). Connectedness, awareness and social presence. *PRESENCE*, pp. 3, 25

Rickel, J. & Johnson, W. (1999). Animated agents for procedural training in virtual reality: Perception, cognition, and motor control. *Applied artificial intelligence*, **13**, 343–382, pp. 153

Riek, L.D., Rabinowitch, T.C., Chakrabarti, B. & Robinson, P. (2009). Empathizing with robots: Fellow feeling along the anthropomorphic spectrum. In *Proceedings of the 3rd International Conference on Affective Computing and Intelligent Interaction and Workshops*, 1–6, Ieee, pp. 28, 29

Riva, G., Davide, F. & IJsselsteijn, W. (2003). *Being There: Concepts, Effects And Measurements of User Presence in Synthetic Environments -emerging Communication*. Emerging Communication Series, I O S Press, Incorporated, pp. 2

Rushforth, M., Gandhe, S., Artstein, R., Roque, A., Ali, S., Whitman, N. & Traum, D. (2009). Varying personality in spoken dialogue with a virtual human. In *Intelligent Virtual Agents*, 541–542, Springer, pp. 153

SAS, C. & O'HARE, G.M.P. (2003). Presence Equation: An Investigation into Cognitive Factors Underlying Presence. *Presence: Teleoperators and Virtual Environments*, **12**, 523–537, pp. 10

SEAH, M.L. & CAIRNS, P. (2008). From immersion to addiction in videogames. In *Proceedings of the 22nd British HCI Group Annual Conference on HCI 2008: People and Computers XXII: Culture, Creativity, Interaction-Volume 1*, 55–63, British Computer Society, pp. 31

SLATER, M. (2009). Place illusion and plausibility can lead to realistic behaviour in immersive virtual environments. *Philosophical Transactions of the Royal Society B: Biological Sciences*, **364**, 3549–3557, pp. xii, 4, 5, 6, 17, 23, 26, 73, 94, 152

STEPTOE, W. & STEED, A. (2008). High-fidelity avatar eye-representation. In *Virtual Reality Conference, 2008. VR'08. IEEE*, 111–114, IEEE, pp. 153

STUCKY, S., SHAW, B. & ARK, W. (2009). Virtual environments overview. Tech. rep., DTIC Document, pp. 36, 37

SWARTOUT, W., GRATCH, J., HILL JR, R., HOVY, E., MARSELLA, S., RICKEL, J., TRAUM, D. *et al.* (2006). Toward virtual humans. *AI Magazine*, **27**, 96, pp. 153

TAN, K., ROBINSON, I., CULBERTSON, B. & APOSTOLOPOULOS, J. (2010). Enabling genuine eye contact and accurate gaze in remote collaboration. In *Multimedia and Expo (ICME), 2010 IEEE International Conference on*, 430–435, IEEE, pp. 45

TAO, J. (2009). *Affective Information Processing*. Springer London, London, pp. 29

Thomas, F. & Johnston, O. (1995). *The illusion of life: Disney Animation*. Hyperion, pp. 28, 30, 31

Todorovic, D. (2006). Geometrical basis of perception of gaze direction. *Vision research*, **46**, 3549–3562, pp. xiv, 45, 47, 48

Tomasello, M. (2008). *Origins of human communication*, vol. 2008. MIT press, pp. 2

Wang, N. & Gratch, J. (2009). Rapport and facial expression. In *Proceedings of the 3rd International Conference on Affective Computing and Intelligent Interaction and Workshops (ACII'09)*, 1–6, Ieee, pp. 29

Wiendl, V., Dorfmüller-Ulhaas, K., Schulz, N. & André, E. (2007). Integrating a virtual agent into the real world: The virtual anatomy assistant ritchie. In *Intelligent Virtual Agents*, 211–224, Springer, pp. 30, 34, 39

Witmer, B.G. & Singer, M.J. (1998). Measuring Presence in Virtual Environments: A Presence Questionnaire. *Presence: Teleoperators and Virtual Environments*, **7**, 225–240, pp. 10, 24, 68, 72, 80

Woods, A. & Rourke, T. (2004). Ghosting in anaglyphic stereoscopic images. In *Proceedings of SPIE*, vol. 5291, 354–365, pp. 128, 129

Zanbaka, C., Ulinski, A., Goolkasian, P. & Hodges, L. (2007). Social responses to virtual humans: implications for future interface design. In *Proceedings of the SIGCHI conference on Human factors in computing systems*, 1561–1570, ACM, pp. 32, 33, 39